

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

```
Run on:      October 27, 2003, 19:11:58 ; Search time 33.6197 Seconds
              (without alignments)
              1902.657 Million cell updates/sec
```

Title: US-10-017-372E-33
Perfect score: 2139
Sequence: 1 MAPSGRLRLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query						
No.	Score	Match	Length	DB	ID	Description	
1	2050.5	95.9	390	23	AAE13596	Porcine transformi	
2	1928.5	90.2	390	22	AAM39186	Human polypeptide	
3	1921.5	89.8	390	7	AAP61468	PreTGF-beta gene p	
4	1921.5	89.8	390	11	AAR04034	Sequence of pre-TG	
5	1921.5	89.8	390	11	AAR05258	Human pre-transfor	
6	1921.5	89.8	390	12	AAR13813	Human pro-TGF-beta	
7	1921.5	89.8	390	16	AAR73596	Human TGF-beta 1 p	
8	1921.5	89.8	390	17	AAR90827	Pre-transforming g	
9	1918.5	89.7	390	23	AAU77101	Human transforming	
10	1918.5	89.7	390	23	AAE16943	Human transforming	
11	1918	89.7	391	24	ABB82780	TGFB1 Arg25Pro pol	
12	1916.5	89.6	390	13	AAR20124	Sequence of simian	
13	1912.5	89.4	390	15	AAR46227	Human pre-TGF-beta	
14	1911	89.3	391	16	AAR83054	Transforming growt	
15	1910.5	89.3	390	19	AAW78785	Human pre-transfor	
16	1909.5	89.3	390	22	AAB84601	Nucleotide sequenc	
17	1907	89.2	391	9	AAP81362	Human transforming	
18	1906	89.1	434	11	AAR03743	Monkey transformin	
19	1902.5	88.9	390	24	ABB82781	TGFB1 Arg25Pro pol	
20	1885.5	88.1	386	11	AAR05663	Simian Transformin	
21	1880	87.9	387	11	AAR05664	Simian Transformin	
22	1869.5	87.4	390	11	AAR05492	Chimeric simian TG	
23	1865.5	87.2	390	13	AAR27522	TGF-beta 1/beta 2	
24	1846	86.3	389	13	AAR29657	TGF-beta 1. Homo	
25	1841	86.1	453	22	AAM40972	Human polypeptide	
26	1760.5	82.3	390	13	AAR20126	Sequence of hybrid	
27	1759.5	82.3	390	11	AAR05749	Human TGF-Beta2 ex	
28	1753.5	82.0	390	11	AAR05665	Human Transforming	
29	1751.5	81.9	390	11	AAR05666	Hybrid transformin	
30	1744	81.5	391	10	AAP91900	Sequence encoded b	
31	1716	80.2	389	16	AAR79921	Simian-human hybri	
32	1300	60.8	278	15	AAR53090	Polypeptide cross-	
33	1295	60.5	278	12	AAR12541	Latency associated	
34	1262.5	59.0	458	23	ABG31507	LAP-mIFNB construc	
35	1262.5	59.0	463	23	ABG31510	LAP-huIFNB constru	
36	1171	54.7	290	22	ABG06792	Novel human diagno	
37	1135	53.1	450	23	ABG31508	mIFNB-LAP construc	
38	1118	52.3	448	23	ABG31509	huIFNB-LAP constru	
39	944	44.1	227	22	ABG20234	Novel human diagno	
40	903	42.2	236	22	ABG20233	Novel human diagno	
41	870.5	40.7	382	21	AAB08338	Amino acid sequenc	
42	870.5	40.7	382	23	AAU77105	Frog transforming	
43	834.5	39.0	456	19	AAW78786	Pig transforming g	
44	833.5	39.0	412	16	AAR73598	Human TGF-beta 3 p	
45	829.5	38.8	412	13	AAR20621	Transforming Growt	

ALIGNMENTS

RESULT 1

AAE13596

ID AAE13596 standard; Protein; 390 AA.

XX

AC AAE13596;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-beta1) mutant.

XX

KW Porcine; transforming growth factor beta 1; TGF-beta1; gene therapy;
KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;
KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;
KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; mutant;
KW mutein.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT Misc-difference 223

FT /note= "Wild type Cys substituted with Ser"

FT Misc-difference 225

FT /note= "Wild type Cys substituted with Ser"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12980.

XX

PR 20-APR-2000; 2000US-199014P.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Strober W, Nakamura K, Kitani A, Fuss IJ;

XX

DR WPI; 2002-026155/03.

DR N-PSDB; AAD22696.

XX

PT Composition for treating autoimmune diseases e.g. inflammatory bowel
PT disease in humans, comprises vector containing transforming growth
PT factor-beta under the control of inducible promoter -

XX

PS Example 1; Fig 1; 78pp; English.

XX

CC The invention relates to a composition containing a vector comprising a
CC gene encoding a regulatory transcription factor under the control of a
CC promoter encoding a transforming growth factor-beta (TGF-beta). The
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2
CC or TGF-beta3, its variants or homologues, by transfecting a cell which
CC is part of a host suspected of having an autoimmune disease, especially
CC inflammatory bowel disease (IBD), under conditions such that the
CC polypeptide encoded by the nucleic acid sequence in the vector is
CC expressed. The vector is delivered using a delivery system. The delivery
CC of the vector results in substantial elimination of symptoms of the
CC autoimmune disease and increased production of IL-10 by the host. The
CC composition is useful for treating various diseases with an autoimmune

CC component such as multiple sclerosis, rheumatoid arthritis, systemic
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis
CC and psoriasis, and also for assaying the expression of a gene in a cell.
CC The vector is further useful for screening of the effect of test
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.
CC The present sequence is porcine TGF-beta1 mutant.

XX

SQ Sequence 390 AA;

Query Match 95.9%; Score 2050.5; DB 23; Length 390;
Best Local Similarity 96.8%; Pred. No. 7.6e-176;
Matches 390; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

```
QY      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAIRGQILSKLRLA 60
      |||
DB      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAIRGQILSKLRLA 60

QY     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||
DB     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120

QY    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      |||
DB    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180

QY    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      |||
DB    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240

QY    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      |||
DB    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRR-----ALDTNYCFS 287

QY    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      |||
DB    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

QY    361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      |||
DB    348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 2

AAM39186

ID AAM39186 standard; Protein; 390 AA.

XX

AC AAM39186;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 2331.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI58342.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 4; SEQ ID NO 2331; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 390 AA;

Query Match 90.2%; Score 1928.5; DB 22; Length 390;
 Best Local Similarity 90.8%; Pred. No. 7.2e-165;
 Matches 366; Conservative 10; Mismatches 14; Indels 13; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Db	1	MPPSGLRLPLLLPLLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKQSTHSIYMPFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRHRRALDTNDYKDDDDKALDTNYCFS	300
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 3

AAP61468

ID AAP61468 standard; Protein; 390 AA.

XX

AC AAP61468;

XX

DT 31-OCT-2002 (updated)

DT 28-OCT-1991 (first entry)

XX

DE PreTGF-beta gene product.

XX

KW Transforming growth factor beta; cancer; wound healing.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT Protein 279..390

XX

PN EP200341-A.

XX

PD 10-DEC-1986.

XX

PF 21-MAR-1986; 86EP-0302112.

XX

PR 22-MAR-1985; 85US-0715142.

PR 13-MAR-1987; 87US-0025423.

XX

PA (GETH) GENENTECH INC.

XX

PI Derynck RMA;
 XX
 DR WPI; 1986-326875/50.
 DR N-PSDB; AAN60972.
 XX
 PT TGF-beta prodn. from transformed hosts - useful esp. for treating
 PT wounds (J6 2/9/86).
 XX
 PS Disclosure; Fig 1b; 26pp; English.
 XX
 CC The gene product is known to stimulate cell proliferation and
 CC inhibit anchorage-dependent growth of a variety of human cancer cell
 CC lines, it is esp. useful in treatment of burns and the promotion of
 CC surface and internal wound healing. TGF-beta may be expressed from a
 CC transformed CHO cell line.
 CC (Updated on 31-OCT-2002 to add missing OS field.)
 XX
 SQ Sequence 390 AA;

Query Match 89.8%; Score 1921.5; DB 7; Length 390;
 Best Local Similarity 90.6%; Pred. No. 3.1e-164;
 Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

Qy	1	MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAI	RGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAI	RGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEV	TRVLMVESGNQI	120
		:		
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEV	TRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQH	VELYQKYSNDSWR	180
		: :		
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQH	VELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSK	DNTLHVEINGFN	240
		:		
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDS	RDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRRHRALDTNDYK	DDDDKALDTNYCFS	300
		:		
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRRHR-----	ALDTNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT	QYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT	QYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403	
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390	

RESULT 4
 AAR04034
 ID AAR04034 standard; protein; 390 AA.
 XX
 AC AAR04034;

```

XX DT 25-MAR-2003 (updated)
XX DT 31-OCT-2002 (updated)
XX DT 31-MAY-1989 (first entry)
XX
DE Sequence of pre-TGF-beta 1.
XX
KW Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth
KW inhibition.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 348..500
XX
PN W08912101-A.
XX
PD 14-DEC-1989.
XX
PF 08-JUN-1988; 88WO-US01945.
XX
PR 08-JUN-1988; 88WO-US01945.
XX
PA (GETH ) GENENTECH INC.
XX
PI Dernyck RMA, Goeddel DV;
XX
DR WPI; 1990-007474/01.
DR N-PSDB; AAQ02815.
XX
PT Nucleotide sequence encoding transforming growth factor beta-3 used as a
PT probe, or to produce TGF beta 3, for inhibiting growth of certain normal
PT and neoplastic cells, eg A549.
XX
PS Disclosure; Fig. 2; 61pp; English.
XX
CC Sequence is an exon of transforming growth factor-beta 1 (pre-TGF-beta
CC 1) polypeptide and corresponds to AA's 288-338 of mature TGF-beta 1. The
CC nucleic acid encoding second subtype of TGF-beta (TGF-beta 3) is useful
CC as a probe or to produce TGF-beta 3 for inhibition of normal and
CC neoplastic cell growth.
CC (Updated on 31-OCT-2002 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 390 AA;

Query Match 89.8%; Score 1921.5; DB 11; Length 390;
Best Local Similarity 90.6%; Pred. No. 3.1e-164;
Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy 61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
   | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


QY 361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
 ||||||||||||||||||||||||||||||||||||||||
 Db 348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 6

AAR13813

ID AAR13813 standard; Protein; 390 AA.

XX

AC AAR13813;

XX

DT 20-NOV-1991 (first entry)

XX

DE Human pro-TGF-beta 1.

XX

KW Osteogenetic; tumoricidal.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..29

FT /note= "signal peptide"

FT Peptide 30..390

FT /note= "pro-TGF-beta 1"

FT Peptide 279..390

FT /note= "TGF-beta 1"

XX

PN JP03180192-A.

XX

PD 06-AUG-1991.

XX

PF 07-DEC-1989; 89JP-0318243.

XX

PR 07-DEC-1989; 89JP-0318243.

XX

PA (KIRI) KIRIN BREWERY KK.

XX

DR WPI; 1991-271579/37.

DR N-PSDB; AAQ13392.

XX

PT Human pro-TGF-beta 1 prodn., for osteo-genetic activity - by

PT preparing DNA chain contg. base sequence coding for human

PT pre:pro-TGF-beta 1, forming expression vector etc.

XX

PS Claim 1; Fig 1; 16pp; Japanese.

XX

CC The amino acid sequence codes for human prepro-TGF-beta 1 which

CC can be produced by recombinant methods, it has osteogenetic and

CC tumoricidal activity.

XX

SQ Sequence 390 AA;

Query Match 89.8%; Score 1921.5; DB 12; Length 390;

Best Local Similarity 90.6%; Pred. No. 3.1e-164;

Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

Qy	1	MAPSGLRLLP LLLP LLWLLV LTPGRPAAGLSTCKTIDMELV KRKR IEAIRGQILSKLR LA	60
Db	1	MPPSGLRLLP LLLP LLWLLV LTPGPPAAGLSTCKTIDMELV KRKR IEAIRGQILSKLR LA	60
Qy	61	SPPSQGDVPPG PLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPG PLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEI EGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATI HGMRNPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS	300
Db	241	TGRRGDLATI HGMRNPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 7

AAR73596

ID AAR73596 standard; Protein; 390 AA.

XX

AC AAR73596;

XX

DT 25-MAR-2003 (updated)

DT 20-DEC-1995 (first entry)

XX

DE Human TGF-beta 1 protein.

XX

KW Transforming growth factor-beta; Human TGF-beta protein; TGF-beta 1;

KW TGF-beta 3; osteogenic cell source; OCS; bone deficiency;

KW bone-inducing cofactor.

XX

OS Homo sapiens.

XX

PN US5409896-A.

XX

PD 25-APR-1995.

XX

PF 12-NOV-1993; 93US-0132405.

XX

PR 01-SEP-1989; 89US-0401906.

PR 12-NOV-1991; 91US-0790856.

PR 18-MAY-1993; 93US-0063841.

PR 12-NOV-1993; 93US-0132405.

XX

RESULT 8

AAR90827

ID AAR90827 standard; Protein; 390 AA.

XX

AC AAR90827;

XX

DT 25-MAR-2003 (updated)

DT 25-JAN-1980 (first entry)

XX

DE Pre-transforming growth factor beta 1.

XX

KW transforming growth factor beta 1; wound healing;
KW recombinant production.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 8..23

FT /note= "hydrophobic domain"

FT Modified-site 82..84

FT /note= "potential N-glycosylation site"

FT Modified-site 136..138

FT /note= "potential N-glycosylation site"

FT Modified-site 176..178

FT /note= "potential N-glycosylation site"

FT Cleavage-site 277..279

FT /note= "trypsin-like peptidase cleavage site"

FT Protein 279..390

FT /label= mature_TGF_beta_1

XX

PN US5482851-A.

XX

PD 09-JAN-1996.

XX

PF 05-NOV-1993; 93US-0147364.

XX

PR 13-MAR-1987; 87US-0025423.

PR 22-MAR-1985; 85US-0715142.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992; 92US-0845893.

PR 05-NOV-1993; 93US-0147364.

XX

PA (GETH) GENENTECH INC.

XX

PI Derynck RMA, Goeddel DV;

XX

DR WPI; 1996-076891/08.

DR N-PSDB; AAT15720.

XX

PT New recombinant human transforming growth factor-beta prods. - produced
PT using Chinese hamster ovary cells, for use in diagnostic applications
PT or in therapy

XX

PS Example 3; Fig 1A-C; 26pp; English.

XX

CC The pre-transforming growth factor (TGF) beta 1 protein is encoded
 CC by AAT15720. The mature TGF beta 1 monomer is cleaved from the
 CC precursor at the Arg-Arg dipeptide immediately preceding the mature
 CC TGF-beta 1 NH2-terminus. It does not contain a recognisable N-terminal
 CC signal peptide typical of most secreted proteins. The pre-TGF beta 1
 CC contains several pairs of basic residues which could undergo
 CC post-translational cleavage and give rise to separate polypeptide
 CC entities. The precursor contains 3 potential N-glycosylation sites, none
 CC of which are localised in the mature TGF beta 1. This is useful in
 CC purification of the mature protein. TGF beta 1 can be used in, e.g. wound
 CC healing.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 390 AA;

Query Match 89.8%; Score 1921.5; DB 17; Length 390;
 Best Local Similarity 90.6%; Pred. No. 3.1e-164;
 Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

QY	1	MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLTTPGPFAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
QY	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		: : :	
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
QY	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		: : :	
Db	121	YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
QY	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
		: :	
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSDSRDNTLQVDINGFT	240
QY	241	SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS	300
		:	
Db	241	TGRRGDLATIHGMRNPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS	287
QY	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
QY	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 9

AAU77101

ID AAU77101 standard; Protein; 390 AA.

XX

AC AAU77101;

XX

DT 05-JUN-2002 (first entry)

XX

DE Human transforming growth factor beta 1 (TFG-beta-1) polypeptide.

XX
 KW Human; transforming growth factor beta; TGF-beta; insulin production;
 KW type I diabetes mellitus; pancreatic cell outgrowth; wound healing;
 KW pancreatic duct tissue; ischaemia; stroke; nervous system aging;
 KW neurological condition; neurodegenerative disease; inflammation;
 KW vasal injury; chemical injury; traumatic injury; tumour-induced injury;
 KW amyotrophic lateral sclerosis; spinocerebellar degeneration;
 KW immunological disease; multiple sclerosis; TGF-beta-1.
 XX
 OS Homo sapiens.
 XX
 PN WO200212336-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 09-FEB-2001; 2001WO-US04192.
 XX
 PR 09-AUG-2000; 2000US-0635368.
 XX
 PA (CURI-) CURIS INC.
 XX
 PI Wang M, Pang K;
 XX
 DR WPI; 2002-257468/30.
 XX
 PT Treating a subject with a disorder resulting from insufficient insulin
 PT production, and inducing outgrowth of pancreatic cells, involves using
 PT a transforming growth factor beta therapeutic -
 XX
 PS Disclosure; Fig 1; 77pp; English.
 XX
 CC The invention relates to treating a subject with a disorder resulting
 CC from insufficient insulin production, involving contacting the subject
 CC with a transforming growth factor beta (TGF-beta) therapeutic. TGF-beta
 CC polypeptides can be used for treating a subject with a disorder resulting
 CC from insufficient insulin production, e.g. type I diabetes mellitus, and
 CC for inducing outgrowth of pancreatic cells associated with pancreatic
 CC duct tissue within a subject. A composition comprising a TGF-beta protein
 CC may be useful in wound healing and treatment of neurological conditions
 CC derived from acute, subacute or chronic injury to the nervous system,
 CC including traumatic injury, chemical injury, vasal injury and deficits
 CC (such as ischaemia resulting from stroke), together with
 CC infectious/inflammatory and tumour-induced injury, aging of the nervous
 CC system including Alzheimer's disease, chronic neurodegenerative diseases
 CC including Parkinson's disease, Huntington's chorea, amyotrophic lateral
 CC sclerosis, spinocerebellar degenerations and chronic immunological
 CC diseases of the nervous system or affecting the nervous system, including
 CC multiple sclerosis. This sequence represents the human TGF-beta-1
 CC protein.
 XX
 SQ Sequence 390 AA;

Query Match 89.7%; Score 1918.5; DB 23; Length 390;
 Best Local Similarity 90.6%; Pred. No. 5.7e-164;
 Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAI RQILSKLRLA 60

Db	1	MPPSGLRLLLLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS	300
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 10

AAE16943

ID AAE16943 standard; Protein; 390 AA.

XX

AC AAE16943;

XX

DT 18-APR-2002 (first entry)

XX

DE Human transforming growth factor-beta1 (TGF-beta1) protein.

XX

KW Human; transforming growth factor-beta1; TGF-beta1; osteoporosis;
 KW latency associated peptide; LAP; integrin alphavbeta3; apoptosis;
 KW immunomodulation; inflammatory disease; fibrotic disease; cancer;
 KW diabetic retinopathy; chronic obstructive pulmonary disorder;
 KW bone resorption; rheumatoid arthritis; psoriasis; restenosis;
 KW atherosclerosis; liver fibrosis; asthma; cytostatic; osteopathic;
 KW ophthalmological; antiarteriosclerotic; vasotropic.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..29

FT /label= Signal_peptide

FT Region 30..278

FT /note= "LAP-beta1"

FT Domain 244..246

FT /note= "RGD motif"

FT Protein 279..390

FT /note= "Human mature TGF-beta1 protein"

Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWL SRGGEIEGFRLSAHCSDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMRNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALD TNYCFS	300
		:	
Db	241	TGRRGDLATIHGMRNRPFLLLMATPLERAQHLQSSRHRR-----ALD TNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 11

ABB82780

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DR

PT

XX

XX

CC renal failure and involves detecting the presence of a genetic
 CC polymorphism pattern in transforming growth factor beta 1 (TGFB1) gene in
 CC a sample from an individual, where polymorphism pattern is associated
 CC with renal failure. The method is useful for determining an individual's
 CC susceptibility to the progression of renal failure. The nucleic acid
 CC comprising a T(-509)C polymorphism of TGFB1 gene, or a polypeptide
 CC comprising a sequence of 391 amino acids is useful for preparing a
 CC medicament for retarding or preventing the progression of renal disease,
 CC and for drug research purposes for retarding or preventing the
 CC progression of renal disease. Sequences ABV75386-88 represents the
 CC protein sequence for the TGFB1 G-allele of the Arg25Pro polymorphism
 CC of exon 1.

XX

SQ Sequence 391 AA;

Query Match 89.7%; Score 1918; DB 24; Length 391;
 Best Local Similarity 90.6%; Pred. No. 6.4e-164;
 Matches 366; Conservative 10; Mismatches 14; Indels 14; Gaps 2;

Qy	1	MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIE	60
Db	1	MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIE	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		: : :	
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLL-RLKLKVEQHVELYQKYSNDSW	179
		: : :	
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRRLKLKVEQHVELYQKYSNNSW	180
Qy	180	RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGF	239
		: :	
Db	181	RYLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGF	240
Qy	240	NSGRRGDLATIHGMNRPFLLLMTPLERAQHLHSSRRHRRALDTNDYKDDDDKALDTNYCF	299
		:	
Db	241	TTGRRGDLATIHGMNRPFLLLMTPLERAQHLQSSRHR-----ALDTNYCF	287
Qy	300	SSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQH	359
Db	288	SSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQH	347
Qy	360	NPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	NPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	391

RESULT 12

AAR20124

ID AAR20124 standard; Protein; 390 AA.

XX

AC AAR20124;

XX

DT 25-MAR-2003 (updated)

DT 16-APR-1992 (first entry)

XX

DE Sequence of simian transforming growth factor (TGF) beta-1.
 XX
 KW Hypertension therapy; hypotensive agent; blood pressure modulator.
 XX
 OS Monkey.
 XX
 FH Key Location/Qualifiers
 FT Peptide 8..21
 FT Protein 279..390
 XX
 PN W09119513-A.
 XX
 PD 26-DEC-1991.
 XX
 PF 20-JUN-1991; 91WO-US04449.
 XX
 PR 20-JUN-1990; 90US-0541221.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Oleson FB, Comereski CR;
 XX
 DR WPI; 1992-024199/03.
 DR N-PSDB; AAQ20289.
 XX
 PT Use of transforming growth factor (TGF)-beta and their
 PT antagonists - for modulating blood pressure, for treating
 PT hypertension and hypotension
 XX
 PS Disclosure; Fig 1; 42pp; English.
 XX
 CC A new method for treating hypertension comprises administering a
 CC transforming growth factor (TGF)-beta to an individual at a dose
 CC effective for lowering blood pressure; the TGF-beta may be e.g.
 CC mature TGF-beta, TGF-beta2, a mature TGF-beta1/beta2 hybrid, TGF-
 CC betal precursor, a latent TGF-beta2 precursor, hybrid TGF-beta1/TGF-
 CC beta2 precursor, a latent TGF-beta1 complex or a latent TGF-beta2
 CC complex.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 390 AA;

Query Match 89.6%; Score 1916.5; DB 13; Length 390;
 Best Local Similarity 90.3%; Pred. No. 8.7e-164;
 Matches 364; Conservative 10; Mismatches 16; Indels 13; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTPSRPAAGLSTCKTIDMELVKRKRIETIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | : | : |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | : | | | | | | | | | | | | | | | | | | | | | : | |
Db    121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

```

```

Qy      181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
        |||||:|||||:| ||||| | ||||| |:|||
Db      181 YLSNRLLAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSKDNTLQVDINGFT 240

Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
        :||||| ||||| ||||| |||||
Db      241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS 287

Qy      301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
        ||||| ||||| ||||| ||||| |||||
Db      288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy      361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
        ||||| ||||| ||||| ||||| |||||
Db      348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 13

AAR46227

ID AAR46227 standard; Protein; 390 AA.

XX

AC AAR46227;

XX

DT 25-MAR-2003 (updated)

DT 09-JUL-1994 (first entry)

XX

DE Human pre-TGF-beta-1.

XX

KW TGF-beta-1; TGF-beta-2; transforming growth factor beta-1;
KW transforming growth factor beta-3; recombinant; wound healing;
KW vulnerary.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	Peptide	279..390
FT		/label= Mat_peptide
FT	Cleavage-site	279
FT		/note= "TGF-beta-1 release site"
FT	Modified-site	82..84
FT		/label= N-glycosylation_site
FT	Modified-site	136..138
FT		/label= N-glycosylation_site
FT	Modified-site	176..178
FT		/label= N-glycosylation_site

XX

PN US5284763-A.

XX

PD 08-FEB-1994.

XX

PF 04-MAR-1992; 92US-0845893.

XX

PR 22-MAR-1985; 85US-0715142.

PR 13-MAR-1987; 87US-0025423.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992; 92US-0845893.

RESULT 14

AAR83054

ID AAR83054 standard; Protein; 391 AA.

XX

AC AAR83054;

XX

DT 25-JUN-1996 (first entry)

XX

DE Transforming growth factor-beta 1.

XX

KW macrophage inducible nitric oxide synthase; iNOS; constitutive NOS;

KW interleukin-1-beta; transforming growth factor-beta; TGF-beta; IL1-beta;

KW nitric oxide production; hypotension; inflammation; septic shock;

KW treatment.

XX

OS Mammalian sp.

XX

FH Key Location/Qualifiers

FT Protein 279..391

FT /note= "represents the mature active TGF beta-1 mol."

XX

PN WO9526745-A1.

XX

PD 12-OCT-1995.

XX

PF 05-APR-1994; 94WO-US03705.

XX

PR 05-APR-1994; 94WO-US03705.

XX

PA (HARD) HARVARD COLLEGE.

XX

PI Lee M, Perrella MA;

XX

DR WPI; 1995-358443/46.

DR N-PSDB; AAT05876.

XX

PT Treatment of hypotension, esp. in septic shock - by administering

PT transforming growth factor-beta e.g. to inhibit inducible nitric

PT oxide synthase gene transcription

XX

PS Disclosure; Fig 17; 52pp; English.

XX

CC Transforming growth factor-beta 1 (TGF-beta 1) has been found to inhibit

CC inducible nitric oxide synthase (iNOS) gene transcription, esp. in

CC interleukin-1-beta (IL1-beta) stimulated rat smooth muscle cells, and at

CC a dose which does not inhibit constitutive NOS. TGF-beta 1 or 2

CC (AAR83055) or their active fragments (esp. derived from the

CC carboxy-terminal 112 amino acids), can be used in the treatment of

CC hypotension, such as that associated with severe inflammation or septic

CC shock.

XX

SQ Sequence 391 AA;

Query Match 89.3%; Score 1911; DB 16; Length 391;

Best Local Similarity 90.3%; Pred. No. 2.7e-163;

Matches 365; Conservative 10; Mismatches 15; Indels 14; Gaps 2;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | : | : |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLL-RLKLKVEQHVELYQKYSNDSW 179
      | | | | | : | | : | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRRLKLKVEQHVELYQKYSNNSW 180

Qy    180 RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGF 239
      | | | | | | | | | | | | | | | | | | | : | | | | | | | | | : | | | |
Db    181 RYLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEI EGFRLSAHCSCDSRDNTLQVDINGF 240

Qy    240 NSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCF 299
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHR-----ALDTNYCF 287

Qy    300 SSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQH 359
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 SSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQH 347

Qy    360 NPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 NPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 391

```

RESULT 15

AAW78785

ID AAW78785 standard; Protein; 390 AA.

XX

AC AAW78785;

XX

DT 25-MAR-2003 (updated)

DT 21-DEC-1998 (first entry)

XX

DE Human pre-transforming growth factor-beta 1.

XX

KW Transforming growth factor-beta 1; TGF-beta 1; human.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 8..23

FT /note= "hydrophobic domain"

FT Protein 279..390

FT /label= Mat_protein

FT Modified-site 82..84

FT /note= "Asn is N-glycosylated"

FT Modified-site 136..138

FT /note= "Asn is N-glycosylated"

FT Modified-site 176..178

FT /note= "Asn is N-glycosylated"

FT Cleavage-site 277..278

FT /note= "cleavage site for relase of TGF-beta 1"

XX

PN US5801231-A.

XX

PD 01-SEP-1998.

XX

PF 30-MAY-1995; 95US-0454468.

XX

PR 13-MAR-1987; 87US-0025423.

PR 22-MAR-1985; 85US-0715142.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992; 92US-0845893.

PR 05-NOV-1993; 93US-0147364.

PR 30-MAY-1995; 95US-0454468.

XX

PA (GETH) GENENTECH INC.

XX

PI Derynck RMA, Goeddel DV;

XX

DR WPI; 1998-494840/42.

DR N-PSDB; AAV52933.

XX

PT DNA encoding transforming growth factor-beta precursor sequence -

PT useful for analysis to perform manipulations to increase yield of

PT recombinant production of the protein

XX

PS Example 3; Fig 1B 1-3; 26pp; English.

XX

CC This is the amino acid sequence of human transforming growth
CC factor-beta 1 precursor (preTGF-beta 1). It was deduced from
CC a preTGF-beta 1 cDNA sequence (see AAV52933). The invention relates
CC to the recombinant production of TGF-beta. Biologically active
CC TGF-beta is defined as being capable of inducing EGF-potentiated
CC anchorage independent growth of target cell lines and/or growth
CC inhibition of neoplastic cell lines. Nucleic acids encoding
CC TGF-beta have been isolated and cloned into vectors which are
CC replicated in bacteria and expressed in eukaryotic cells. TGF-beta
CC recovered from transformed cells is used in known therapeutic
CC applications.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 390 AA;

Query Match 89.3%; Score 1910.5; DB 19; Length 390;
Best Local Similarity 90.1%; Pred. No. 3e-163;
Matches 363; Conservative 10; Mismatches 17; Indels 13; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Db 1 MPPSGLRLLPLLLPLLWLLVLTPGPPAPGLSTCKTIDMEQVKRKRIEAIRGQILSKLRLA 60

Qy 61 SPFSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120

Db 61 SPFSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180

```

Db      121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLRLKLVKVEQHVELYQKYSNNSWR 180
Qy      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
        |||:|||||
Db      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240
Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
        :|||:|||||
Db      241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS 287
Qy      301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
        |||:|||||
Db      288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347
Qy      361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
        |||:|||||
Db      348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

Search completed: October 28, 2003, 09:06:48
Job time : 34.6197 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 09:09:54 ; Search time 24.4507 Seconds
(without alignments)
2760.110 Million cell updates/sec

Title: US-10-017-372E-33
Perfect score: 2139
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%	Query				
No.	Score	Match	Length	DB	ID	Description	
1	2032.5	95.0	390	11	US-09-214-592-26	Sequence 26, Appl	
2	1931.5	90.3	390	11	US-09-214-592-29	Sequence 29, Appl	
3	1928.5	90.2	390	15	US-10-087-268-2	Sequence 2, Appli	
4	1921.5	89.8	390	15	US-10-087-268-5	Sequence 5, Appli	
5	1918.5	89.7	390	12	US-10-276-947-1	Sequence 1, Appli	
6	1916.5	89.6	390	11	US-09-214-592-33	Sequence 33, Appl	
7	1911	89.3	391	11	US-09-214-592-17	Sequence 17, Appl	
8	1907.5	89.2	390	10	US-09-756-283A-23	Sequence 23, Appl	
9	1895.5	88.6	390	11	US-09-214-592-28	Sequence 28, Appl	
10	1837.5	85.9	390	11	US-09-214-592-20	Sequence 20, Appl	
11	1837.5	85.9	390	11	US-09-214-592-23	Sequence 23, Appl	
12	1574.5	73.6	315	11	US-09-214-592-25	Sequence 25, Appl	
13	1262.5	59.0	455	10	US-09-756-283A-20	Sequence 20, Appl	
14	1139	53.2	447	10	US-09-756-283A-22	Sequence 22, Appl	
15	1049	49.0	373	11	US-09-214-592-32	Sequence 32, Appl	
16	873.5	40.8	412	11	US-09-214-592-31	Sequence 31, Appl	
17	870.5	40.7	382	11	US-09-214-592-34	Sequence 34, Appl	
18	870	40.7	383	10	US-09-756-283A-27	Sequence 27, Appl	
19	836	39.1	409	11	US-09-214-592-27	Sequence 27, Appl	
20	832.5	38.9	410	11	US-09-214-592-22	Sequence 22, Appl	
21	831.5	38.9	412	11	US-09-214-592-24	Sequence 24, Appl	
22	829.5	38.8	412	11	US-09-214-592-19	Sequence 19, Appl	
23	829.5	38.8	412	14	US-10-028-158-21	Sequence 21, Appl	
24	829	38.8	414	11	US-09-214-592-21	Sequence 21, Appl	
25	825.5	38.6	412	10	US-09-756-283A-25	Sequence 25, Appl	
26	814	38.1	414	10	US-09-756-283A-24	Sequence 24, Appl	
27	814	38.1	414	11	US-09-214-592-18	Sequence 18, Appl	
28	813	38.0	412	11	US-09-214-592-30	Sequence 30, Appl	
29	794	37.1	304	10	US-09-756-283A-26	Sequence 26, Appl	
30	755.5	35.3	139	14	US-10-002-278-8	Sequence 8, Appli	
31	640	29.9	114	10	US-09-813-459-22	Sequence 22, Appl	
32	640	29.9	114	14	US-10-115-406-21	Sequence 21, Appl	
33	640	29.9	114	15	US-10-154-333-23	Sequence 23, Appl	
34	640	29.9	115	10	US-09-859-211-47	Sequence 47, Appl	
35	640	29.9	115	10	US-09-880-708-25	Sequence 25, Appl	

36	640	29.9	115	11	US-09-872-856-47	Sequence 47, Appl
37	640	29.9	115	15	US-10-335-483-29	Sequence 29, Appl
38	638	29.8	112	10	US-09-813-271B-2	Sequence 2, Appli
39	638	29.8	113	10	US-09-813-398-13	Sequence 13, Appl
40	562	26.3	98	12	US-10-187-394-1	Sequence 1, Appli
41	542	25.3	116	14	US-10-115-406-24	Sequence 24, Appl
42	542	25.3	116	15	US-10-154-333-26	Sequence 26, Appl
43	535	25.0	112	10	US-09-813-271B-8	Sequence 8, Appli
44	505	23.6	114	14	US-10-115-406-25	Sequence 25, Appl
45	505	23.6	114	15	US-10-154-333-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1

US-09-214-592-26

; Sequence 26, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki,CMotoo

; APPLICANT: Shibata,CKenji

; APPLICANT: Sato,CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

; FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 26

; LENGTH: 390

; TYPE: PRT

; ORGANISM: porcine

US-09-214-592-26

Query Match 95.0%; Score 2032.5; DB 11; Length 390;

Best Local Similarity 95.8%; Pred. No. 1.4e-181;

Matches 386; Conservative 1; Mismatches 3; Indels 13; Gaps 1;

Qy	1	MAPSGLRLLPLLLPLLWLLVLT	PGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLT	PGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMLVESGNQI	120
Db	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMLVESGNQI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEP	VLLSRAELRLLRLKLKVEQHVELYQKYS	NDNDSWR	180
Db	121	YDKFKGTPHSLYMLFNTSELREAVPEP	VLLSRAELRLLRLKLKVEQHVELYQKYS	NDNDSWR	180
Qy	181	YLSNRLAPSDSPEWLSFDVTVGVVRQWL	TRREAIEGFRLSAHSSSDSKDNTLHVE	INGFN	240
Db	181	YLSNRLAPSDSPEWLSFDVTVGVVRQWL	TRREAIEGFRLSAHSCSDSKDNTLHVE	INGFN	240

Qy	241	SGRRGDLATIHGMRPFLLLMATPLERAQHLHSSRRRALDNDYKDDDDKALDNYCFS	300
Db	241	SGRRGDLATIHGMRPFLLLMATPLERAQHLHSSRRRR-----ALDNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 2

```

US-09-214-592-29
; Sequence 29, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
;   APPLICANT: Yamasaki,CMotoo
;   APPLICANT: Shibata,CKenji
;   APPLICANT: Sato,CYasufumi
;   TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
;   TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
;   FILE REFERENCE: 11060
;   CURRENT APPLICATION NUMBER: US/09/214,592A
;   CURRENT FILING DATE: 1999-01-18
;   NUMBER OF SEQ ID NOS: 34
;   SOFTWARE:
; SEQ ID NO 29
;   LENGTH: 390
;   TYPE: PRT
;   ORGANISM: ovine
US-09-214-592-29

```

Query Match 90.3%; Score 1931.5; DB 11; Length 390;
Best Local Similarity 90.6%; Pred. No. 4e-172;
Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

[illegible]

[illegible]

RESULT 3

US-10-087-268-2

; Sequence 2, Application US/10087268

; Publication No. US20030119010A1

; GENERAL INFORMATION:

: APPLICANT: Jonsonn, Julie Ruth

; APPLICANT: Powell, Elizabeth Ellen

; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or condition

; FILE REFERENCE: Fibrosis

; CURRENT APPLICATION NUMBER: US/10/087,268

CURRENT FILING DATE: 2002-03-01

; NUMBER OF SEQ ID NOS: 6

```

; SOFTWARE: PatentIn version 3.1

```

; SEO ID NO 2

LENGTH: 390

; TYPE: PRT

; ORGANISM: Human

US-10-087-268-2

Query Match 90.2%; Score 1928.5; DB 15; Length 390;

Best Local Similarity 90.8%; Pred. No. 7.6e-172;

Matches 366; Conservative 10; Mismatches 14; Indels 13; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA 60
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1 MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAI R GQILSKLRLA 60

Qy 61 SPSPQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
| | | | | : | | | | | | | | | | | | | | | | | | | | | : | | |

Db 61 SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
| | | | : | : | | | | | | | | | | | | | | | | | | | | | | : |

Db 121 YDKFKOSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEOHVELYOKYSNNSWR 180

QY 181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
 | | | | | | | | | | | | : | | | | | | | | | | : | | | | | |

Db 181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLOVDINGFT 240

QY 241 SGRRGDLATIHGMRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDKALDNYCFS 300
 :| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLOSSRHR-----ALDTNYCFS 287

Qy 301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360

Db 288 STEKNCCVROLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTOYSKVLALYNOHN 347

US-10-087-268-5

Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

Qy 361 PGASAAPCCVPOALEPLPIVYYVGRKPKVEOLSNMIVRSCKCS 403

RESULT 6

US-09-214-592-33

; Sequence 33, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki,CMotoo

; APPLICANT: Shibata,CKenji

; APPLICANT: Sato,CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

; FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 33

; LENGTH: 390

; TYPE: PRT

; ORGANISM: simian

US-09-214-592-33

Query Match 89.6%; Score 1916.5; DB 11; Length 390;

Best Local Similarity 90.3%; Pred. No. 1e-170;

Matches 364; Conservative 10; Mismatches 16; Indels 13; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTPSRPAAGLSTCKTIDMELVKRKRIETIRGQILSKRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | : | : |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | : | | | | | | | | | | | | | | | | | | | | | | : | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | : | | | | | | | | | | : | | | | | | | | | | | : | | |
Db    181 YLSNRLLAPSNSPEWLSFDVTGVVRQWLVRGGEIEGFRLSAHCSCDSKDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 7


```

; Patent No. US20020151478A1
; GENERAL INFORMATION:
;   APPLICANT: Chernajovsky, Yuti
;   APPLICANT: Dreja, Hanna Stina
;   APPLICANT: Adams, Gillian
;   TITLE OF INVENTION: Latent Fusion Protein
;   FILE REFERENCE: 0623.1000000
;   CURRENT APPLICATION NUMBER: US/09/756,283A
;   CURRENT FILING DATE: 2001-01-09
;   NUMBER OF SEQ ID NOS: 100
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
;   LENGTH: 390
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-756-283A-23

```

Query Match 89.2%; Score 1907.5; DB 10; Length 390;
Best Local Similarity 90.1%; Pred. No. 7e-170;
Matches 363; Conservative 10; Mismatches 17; Indels 13; Gaps 1;

Qy		1	MPSGLRLLLPLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
		:		:
Db		1	MPPSGRLRLPLLPLWLLVLTPGPAAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRILA	60
Qy		61	SPPSQGDVPVGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:		:
Db		61	SPPSQGEVPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHHEI	120
Qy		121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:		:
Db		121	YDKFKQSTHSTYMFFNISEREAPEPVLLSRAELRLLRLKLKVEQHVELYQKYSSNSWR	180
Qy		181	YLSNRLLAPSDSPDWLFDFVTGVVRQWLRTREAIEGFRLSAHSSESDKDNTLHVINGFN	240
		:		: :
Db		181	YLSNRLLAPSDSPDWLFDFVTGVVRQLSRGGEGFGFRSLAHSCDSDRDNLQVDINGFT	240
Qy		241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS	300
	:			:
Db		241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSRRHR-----ALDTNYCFS	287
Qy		301	STEKNCCVRQLYIDFRKDLGWKWIIHEPKGYHANFCLGPCPYIWISLDTOYSKVLYALYNQHN	360
		:		:
Db		288	STEKNCCVRQLYIDFRKDLGWKWIIHEPKGYHANFCLGPCPYIWISLDTOYSKVLYALYNQHN	347
Qy		361	PGASAAPCCVPQALEPLPIVVYVGRKPKEQLSNMIIVRSCKCS	403
		:		:
Db		348	PGASAAPCCVPQALEPLPIVVYVGRKPKEQLSNMIIVRSCKCS	390

RESULT 9

US-09-214-592-28

; Sequence 28, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki, CMotoo

; APPLICANT: Shibata, CKenji


```

; APPLICANT: Sato,CYasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 28
; LENGTH: 390
; TYPE: PRT
; ORGANISM: canine
US-09-214-592-28

```

Qy	1	MAPSGRLRLPLLLPLLWLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
		:	
Db	1	MPPSGRLRLPLLLPLLRLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLS	60
Qy	61	SPPSQGDVPPGPGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		: :	
Db	61	SPPSQGEVPVPVPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVENTNKI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKEQHVELYQKYSNDSWR	180
		: : :	
Db	121	YEKVKKSPHSIYMLFNTSELREAVPEPVLLSRAELRLLRLKLKAEQHVELYQKYSNDSWR	180
Qy	181	YLSNRLLAPSDSPEWL SFDVTGVVRQWL TRREA IEGFRLSAHSSSDSKDNTLHVEINGFN	240
		: : : : :	
Db	181	YLSNRLLAPSDTPEWL SFDVTGVVRQWLSHGGEVEGFRLSAHCSCDSKDNTLQVDINGFS	240
Qy	241	SGRRGDLATI HGMNRPFL LLMATPLERA QHLHSSRRHRRALDTNDYKDDDDKALDTNYCFS	300
Db	241	SSRRGDLATI HGMNRPFL LLMATPLERA QHLHSSRQRR-----ALDTNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 20
; LENGTH: 390
; TYPE: PRT
; ORGANISM: murine
US-09-214-592-20

Query Match 85.9%; Score 1837.5; DB 11; Length 390;
Best Local Similarity 85.9%; Pred. No. 2.5e-163;
Matches 346; Conservative 15; Mismatches 29; Indels 13; Gaps 1;

```
Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAI RQILSKRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLPWL LVTTPGRPAAGLSTCKTIDMELVKKRIEAI RQILSKRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | : | |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | : | | | | | : | | | | | : | | | | | | | | | | | | | | | | | : | |
Db    121 YEKTKDISHSIYMFNTSDIREAVPEPPLLSRAELRLQRLKSSVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | : | : | | | | | | | | | | | : : | : | | | | | | | | | :
Db    181 YLGNRL LTPTDTPEWLSFDVTGVVRQWLNQGDGIQGF RFSAHCS CDSKDNKLHVEINGIS 240

Qy    241 SGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRRHRRALDTNDYKDDDDKALDTNYCFS 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLGTI HDMNRPFLLLMATPLERAQHLHSSRRHR-----ALDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 11

US-09-214-592-23

; Sequence 23, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki, CMotoo
; APPLICANT: Shibata, CKenji
; APPLICANT: Sato, CYasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 23
; LENGTH: 390
; TYPE: PRT
; ORGANISM: rat
US-09-214-592-23

Query Match 85.9%; Score 1837.5; DB 11; Length 390;
Best Local Similarity 85.9%; Pred. No. 2.5e-163;
Matches 346; Conservative 14; Mismatches 30; Indels 13; Gaps 1;

```
Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLPWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGGLPEAVLALYNSTRDRVAGESADPEPEPEPEADYYAKEVTRVLMVDRNNAI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | | | | | : | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKTKDITHSIYMFNTSDIREAVPEPPLLSRAELRLQRFKSTVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLGNRLLTPTDTPWLSFDVTGVVRQWLNQGDGIQGFRFSAHCSCDSKDNVLHVEINGIS 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLGTIHDNRPFLLLMATPLERAQHLHSSRHRR-----ALDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 12

US-09-214-592-25

; Sequence 25, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki, CMotoo

; APPLICANT: Shibata, CKenji

; APPLICANT: Sato, CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

; FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 25
; LENGTH: 315
; TYPE: PRT
; ORGANISM: bovine
US-09-214-592-25

Query Match 73.6%; Score 1574.5; DB 11; Length 315;
Best Local Similarity 89.9%; Pred. No. 7.4e-139;
Matches 295; Conservative 9; Mismatches 11; Indels 13; Gaps 1;

```
Qy      76 AVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF 135
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 AILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMF 60

Qy     136 NTSELREAVPEPVLLSRAELRLRLKLVKVEQHVELYQKYSNDSWRYLNSRLLAPSDSPEW 195
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61 NTSELREAVPEPVLLSRADVRLRLKLVKVEQHVELYQKYSNNSWRYLNSRLLAPSDSPEW 120

Qy     196 LSFDTVGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFNSGRRGDLATIHGMNR 255
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db     121 LSFDTVGVVRQWLTRREEIEGFRLSAHCSCDSKDNTLQVDINGFSSGRRGDLATIHGMNR 180

Qy     256 PFLLLMATPLERAQHLHSSSRHRRALDTNDYKDDDDKALDTNYCFSSSTEKNCCVRQLYIDF 315
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db     181 PFLLLMATPLERAQHLHSSSRHRR-----ALDTNYCFSSSTEKNCCVRQLYIDF 227

Qy     316 RKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALE 375
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db     228 RKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALE 287

Qy     376 PLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db     288 PLPIVYYVGRKPKVEQLSNMIVRSCKCS 315
```

RESULT 13
US-09-756-283A-20
; Sequence 20, Application US/09756283A
; Patent No. US20020151478A1
; GENERAL INFORMATION:
; APPLICANT: Chernajovsky, Yuti
; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

;   FEATURE:
;   OTHER INFORMATION: LAP-mIFN    construct
US-09-756-283A-20

```

Query Match 59.0%; Score 1262.5; DB 10; Length 455;
Best Local Similarity 90.5%; Pred. No. 2e-109;
Matches 248; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

[illegible]

```

RESULT 14
US-09-756-283A-22
; Sequence 22, Application US/09756283A
; Patent No. US20020151478A1
; GENERAL INFORMATION:
; APPLICANT: Chernajovsky, Yuti
; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 447
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: mIFN -LAP construct
US-09-756-283A-22

```

Query Match 53.2%; Score 1139; DB 10; Length 447;
Best Local Similarity 87.0%; Pred. No. 7e-98;
Matches 228; Conservative 10; Mismatches 18; Indels 6; Gaps 3;

Qy 14 PL-LWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGDVPPGP 72
 | | | | | | | | | | | | | | | | | | | | | : | | |

Db	190	PLGLW---- <td>245</td>	245
Qy	73	LPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVL MVESGNQIYDKFKGTPHSLY	132
		: : : :	
Db	246	LPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVL MVETHNEIYDKFKQSTHSIY	305
Qy	133	MLFNTSELREAVPEPVLLSRAELRLL-RLKLKVEQHVELYQKYSNDSWRYLSNRLLAPSD	191
		:	
Db	306	MFFNTSELREAVPEPVLLSRAELRLLRRLKLKVEQHVELYQKYSNNSWRYLSNRLLAPSD	365
Qy	192	SPEWL SFDVTGVVRQWL TRREAIEGFRLSAHSSSDSKDNTLHVEINGFN SGRRGDLATIH	251
		: : :	
Db	366	SPEWL SFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFTTGRRGDLATIH	425
Qy	252	GMNRPFLLLMATPLERAQHLHS	273
Db	426	GMNRPFLLLMATPLERAQHLOS	447

US-09-214-592-32

; Sequence 32, Application US/09214592A

; GENERAL INFORMATION:

; APPLICANT: Shibata,CKenji

; APPLICANT: Sato, CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD

FILE REFERENCE: 11060

CURRENT FILING DATE: 1999-01-18

; SOFTWARE:

; SEQ ID NO 32

LENGTH: 373

; TYPE: PRT

; ORGANISM: chicken

Query Match 49.0%; Score 1049; DB 11; Length 373;

Best Local Similarity 53.5%; Pred. No. 1.4e-89;

Qy 30 LSTCKTIDMELVKRKRIEAI RGOILSKLRLASPPSGDVPVPGPLPEAVIALYNSTRDRVA 89

[illegible]

Qv 90 GES-VEPEPEPEADYYAKEVTRVLMVESGNOIYDKFKGTPHSLYMLNTSEI REAVPEPV 148

27 00 022 VEREINIGTES KÖNIGREICH VON GROßBRITANNIEN UND IRLAND 148
: : | : : ||||| : | : : : ||| : |

Qv 149 LLSRAELRLRLRLKLIK-----VEOHVELYOKYSNDSWRYLINRIILAPSDSPEWLSQDVTGV 203

Qy 149 LLSRAELRLLRLKLIK-----VEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGV 203

[illegible]

```

Qy      204 VRQWLTRREAIEGFRLSAHSSSD--SKDNTLHVEINGFNSGRRGDLATIHGMNR--PFL 258
      | |||: | : || | : : || :|||: | : | ||:
Db      181 VHQWLSGSELLGVFKLSVHCPCEMGPGHAEEMRISIEGFEQ-QRGDMQSIKKHRRVPYV 239

Qy      259 LLMATPLERAQHLHSSRRRRALDTNDYKDDDDKALDTNYCF--SSTEKNCCVRQLYIDFR 316
      | || | ||| |||:| || | |||: ||| : ||||| |||||
Db      240 LAMALPAERANELHSARRRR-----DLDTDYCFGPGTDEKNCCVRPLYIDFR 286

Qy      317 KDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAPCCVPQALEP 376
      ||| ||||| ||||| |||||: ||||| ||||| ||||| ||||| |||||
Db      287 KDLQWKWIHEPKGYMANFCMGPCPYIWSADTQYTKVLALYNQHNPASAAPCCVPQTLDP 346

Qy      377 LPIVYYYVGRKPKVEQLSNMIVRSCKCS 403
      |||: |||| : |||||: ||: ||||
Db      347 LPIIYYVGRNVRVEQLSNMVVRACKCS 373

```

Search completed: October 28, 2003, 09:28:02
Job time : 25.4507 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 00:37:16 ; Search time 13.3169 Seconds
(without alignments)
2910.285 Million cell updates/sec

Title: US-10-017-372E-33
Perfect score: 2139
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	2035.5	95.2	390	2	A27512	transforming growth	
2	2004	93.7	391	2	S01413	transforming growth	
3	1931.5	90.3	390	2	I46463	transforming growth	
4	1918.5	89.7	390	1	WFHU2	transforming growth	
5	1916.5	89.6	390	2	A26960	transforming growth	
6	1895.5	88.6	390	2	JC4023	transforming growth	
7	1837.5	85.9	390	1	WFMS2	transforming growth	
8	1837.5	85.9	390	2	S10219	transforming growth	
9	1574.5	73.6	315	2	A40057	transforming growth	
10	1049	49.0	373	2	A41918	transforming growth	
11	873.5	40.8	412	2	A34939	transforming growth	
12	870.5	40.7	382	2	B61036	transforming growth	
13	836	39.1	409	2	S01825	transforming growth	
14	832.5	38.9	410	2	A41397	transforming growth	
15	831.5	38.9	410	2	A55706	transforming growth	
16	829.5	38.8	412	2	A36169	transforming growth	
17	829	38.8	414	1	WFMSB2	transforming growth	
18	814	38.1	414	1	WFMKB2	transforming growth	
19	814	38.1	414	2	A31249	transforming growth	
20	813.5	38.0	413	1	WFXLB2	transforming growth	
21	813	38.0	412	2	A39489	transforming growth	
22	798	37.3	442	2	B31249	transforming growth	
23	693.5	32.4	130	2	I48196	transforming growth	
24	482	22.5	112	2	A61439	transforming growth	
25	279	13.0	425	2	I47072	inhibin beta-A cha	
26	276.5	12.9	424	1	WFPGBA	inhibin beta-A cha	
27	274.5	12.8	424	1	S31440	inhibin beta-A cha	
28	274	12.8	425	1	S50898	inhibin beta-A cha	
29	272.5	12.7	426	1	B24248	inhibin beta-A cha	
30	270.5	12.6	424	1	B40905	inhibin beta-A cha	
31	260.5	12.2	398	2	JH0688	bone morphogenetic	
32	259	12.1	398	2	JH0687	bone morphogenetic	
33	257.5	12.0	413	2	JC4862	activin beta-A cha	
34	256	12.0	394	2	S45355	bone morphogenetic	
35	255	11.9	396	1	BMHU2	bone morphogenetic	
36	253.5	11.9	513	1	BMHU6	bone morphogenetic	
37	247.5	11.6	393	2	S37073	bone morphogenetic	
38	247	11.5	510	2	A54798	Vg-1-related prote	
39	240	11.2	350	2	JC5241	activin beta E cha	
40	237	11.1	455	2	A43918	TGF-beta-related p	
41	233.5	10.9	402	2	A45056	osteogenic protein	
42	233	10.9	420	2	I49541	bone morphogenetic	
43	232	10.8	454	1	BMHU5	bone morphogenetic	
44	229.5	10.7	461	2	S52408	SPDVR1 protein - s	
45	227	10.6	367	2	JC4151	activin beta D cha	

ALIGNMENTS

RESULT 1

A27512
transforming growth factor beta-1 precursor - pig
N;Alternate names: TGF-beta
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 16-Jul-1999
C;Accession: A27512; A26356; I46657
R;Derynck, R.; Rhee, L.
Nucleic Acids Res. 15, 3187, 1987
A;Title: Sequence of the porcine transforming growth factor-beta precursor.
A;Reference number: A27512; MUID:87174844; PMID:3470708
A;Accession: A27512
A;Molecule type: mRNA
A;Residues: 1-390 <DER>
R;Cheifetz, S.; Weatherbee, J.A.; Tsang, M.L.S.; Anderson, J.K.; Mole, J.E.;
Lucas, R.; Massague, J.
Cell 48, 409-415, 1987
A;Title: The transforming growth factor-beta system, a complex pattern of cross-
reactive ligands and receptors.
A;Reference number: A90890; MUID:87102890; PMID:2879635
A;Accession: A26356
A;Molecule type: protein
A;Residues: 279-322 <CHE>
R;Kondaiah, P.; Van Obberghen-Schilling, E.; Ludwig, R.L.; Dhar, R.; Sporn,
M.B.; Roberts, A.B.
J. Biol. Chem. 263, 18313-18317, 1988
A;Title: cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.
Evidence for alternate splicing and polyadenylation.
A;Reference number: I46657; MUID:89054010; PMID:2461367
A;Accession: I46657
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-390 <KON>
A;Cross-references: GB:M23703; NID:g755044; PIDN:AAA64616.1; PID:g755045
C;Genetics:
A;Gene: TGFB; TGF-beta-1
C;Superfamily: inhibin
C;Keywords: growth factor

Query Match 95.2%; Score 2035.5; DB 2; Length 390;
Best Local Similarity 96.0%; Pred. No. 1.2e-155;
Matches 387; Conservative 0; Mismatches 3; Indels 13; Gaps 1;

Qy	1	MAPSGLRLLPLLLPLLWLLVLT	PGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLT	PGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMMVESGNQI	120
Db	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMMVESGNQI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAV	PEPVLLSRAELRLLRLKLKVEQH	VELYQKYSNDSWR	180
Db	121	YDKFKGTPHSLYMLFNTSELREAV	PEPVLLSRAELRLLRLKLKVEQH	VELYQKYSNDSWR	180
Qy	181	YLSNRL LAPSDSPEWLSFDVTG	VVRQWLTRREAIEGFRLSAHSSSD	SKDNTLHVEINGFN	240
Db	181	YLSNRL LAPSDSPEWLSFDVTG	VVRQWLTRREAIEGFRLSAHCSD	SKDNTLHVEINGFN	240

Qy	241	SGRRGDLATIHGMRPFLLLMATPLERAQHLHSSRRRALDNDYKDDDDKALDNYCFS	300
Db	241	SGRRGDLATIHGMRPFLLLMATPLERAQHLHSSRRRR-----ALDNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 2

S01413

transforming growth factor beta-1 precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text change 24-Nov-1999

C;Accession: S01413

R; Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.

Nucleic Acids Res. 16, 8730, 1988

A;Title: Nucleotide sequence of chicken transforming growth factor-beta 1 (TGF-beta 1).

A;Reference number: S01413; MUID:88335639; PMID:3166520

A;Accession: S01413

A;Molecule type: DNA

A;Residues: 1-391 <JAK>

A; Cross-references: EMBL:X12373; NID:q63808; PIDN:CAA30933.1; PID:q63809

C;Superfamily: inhibin

C; Keywords: growth factor

Query Match 93.7%; Score 2004; DB 2; Length 391;
Best Local Similarity 95.0%; Pred. No. 4e-153;
Matches 384; Conservative 0; Mismatches 6; Indels 14; Gaps 2;

Qy	1	M A P S G L R L L P L L L P L L W L L V L T P G R P A A G L S T C K T I D M E L V K R K R I E A I R G Q I L S K L R L A	60
Db	1	M P P S G P G L L P L L L P L L W L L V L T P G R P A A G L S T C K T I D M E L V K R K R I E A I R G Q I L S K L R L A	60
Qy	61	S P P S Q G D V P P G P L P E A V L A L Y N S T R D R V A G E S V E P E P E P E A D Y Y A K E V T R V L M V E S G N Q I	120
Db	61	S P P S Q G D V P P G P L P E A V L A L Y N S T R D R V A G E S V E P E P E P E A D Y Y A K E V T R V L M V E S G N Q I	120
Qy	121	Y D K F K G T P H S L Y M L F N T S E L R E A V P E P V L L S R A E L R L L R L K L K V E Q H V E L Y Q K Y S N D S W R	180
Db	121	Y D K F K G T P H S L Y M L F N T S E L R E A V P E P V L L S R A E L R L L R L K L K V E Q H V E L Y Q K Y S N D S W G	180
Qy	181	Y L S N R L L A P S D S P E W L S F D V T G V V R Q W L T R R E A I E G F R L S A H S S D S K D N T L H V E I N - G F	239
Db	181	Y L S N R L L A P S D S P E W L S F D V T G V V R Q W L T R R E A I E G F R L S A H C S C D S K D N T L H V E I N A G F	240
Qy	240	N S G R R G D L A T I H G M N R P F L L L M A T P L E R A Q H L H S S R H R R A L D T N D Y K D D D K A L D T N Y C F	299
Db	241	N S G R R G D L A T I H G M N R P F L L L M A T P L E R A Q H L H S S R H R R ----- A L D T N Y C F	287
Qy	300	S S T E K N C C V R Q L Y I D F R K D L G W K W I H E P K G Y H A N F C L G P C P Y I W S L D T O Y S K V L A L Y N Q H	359

Db

348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 4

WFHU2

transforming growth factor beta-1 precursor [validated] - human

N;Alternate names: growth-inhibitory factor; TGF type 2; TGF-beta

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1986 #sequence_revision 19-Oct-1995 #text_change 08-Dec-2000

C;Accession: A27513; A01395; A22290; I59664; S53444

R;Derynck, R.; Rhee, L.; Chen, E.Y.; Van Tilburg, A.

Nucleic Acids Res. 15, 3188-3189, 1987

A;Title: Intron-exon structure of the human transforming growth factor-beta precursor gene.

A;Reference number: A27513; MUID:87174845; PMID:3470709

A;Accession: A27513

A;Molecule type: DNA

A;Residues: 1-390 <DER>

A;Cross-references: GB:X05839; GB:Y00112; NID:g37097; PIDN:CAA29283.1;

PID:g1212989

R;Derynck, R.; Jarrett, J.A.; Chen, E.Y.; Eaton, D.H.; Bell, J.R.; Assoian, R.K.; Roberts, A.B.; Sporn, M.B.; Goeddel, D.V.

Nature 316, 701-705, 1985

A;Title: Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells.

A;Reference number: A01395; MUID:85296301; PMID:3861940

A;Accession: A01395

A;Molecule type: mRNA

A;Residues: 1-9, 'P', 11-24, 'P', 26-159, 'R', 160-390 <DE2>

A;Cross-references: GB:X02812; GB:J05114; NID:g37092; PIDN:CAA26580.1;

PID:g37093

A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide

R;Massague, J.; Like, B.

J. Biol. Chem. 260, 2636-2645, 1985

A;Title: Cellular receptors for type beta transforming growth factor. Ligand binding and affinity labeling in human and rodent cell lines.

A;Reference number: A22290; MUID:85131019; PMID:2982829

A;Accession: A22290

A;Molecule type: protein

A;Residues: 279-295, 'XX', 298-301 <MAS>

R;Urushizaki, Y.; Niitsu, Y.; Terui, T.; Koshida, Y.; Mahara, K.; Kohgo, Y.; Urushizaki, I.; Takahashi, Y.; Ito, H.

Tumor Res. 22, 41-55, 1987

A;Title: Cloning and expression of the gene for human transforming growth factor-beta in Escherichia coli.

A;Reference number: I59664

A;Accession: I59664

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 279-390 <RES>

A;Cross-references: GB:M38449; NID:g339557; PIDN:AAA36735.1; PID:g339558

R;Stam, K.; Stewart, A.A.; Qu, G.Y.; Iwata, K.K.; Fenyo, D.; Chait, B.T.; Marshak, D.R.; Haley, J.D.

Biochem. J. 305, 87-92, 1995

A;Title: Physical and biological characterization of a growth-inhibitory activity purified from the neuroepithelioma cell line A673.

A;Reference number: S53444; MUID:95126934; PMID:7826358
A;Accession: S53444
A;Status: preliminary
A;Molecule type: protein
A;Residues: 279-297 <STA>
C;Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide; the active molecule is a dimer of identical polypeptide chains linked by an interchain disulfide bond.
C;Genetics:
A;Gene: GDB:TGFB1; TGFB
A;Cross-references: GDB:120729; OMIM:190180
A;Map position: 19q13.2-19q13.2
C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor; homodimer; mitogen; transformation
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-278/Domain: propeptide #status predicted <PRO>
F;244-246/Region: cell attachment (R-G-D) motif
F;279-390/Product: transforming growth factor beta-1 #status experimental <MAT>
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.7%; Score 1918.5; DB 1; Length 390;
Best Local Similarity 90.6%; Pred. No. 2.9e-146;
Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLLLLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | : | | | | | | | | | | | | | | | |
Db    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 5
A26960
transforming growth factor beta-1 precursor - green monkey
C;Species: Cercopithecus aethiops (green monkey, grivet)

C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 24-Nov-1999
C;Accession: A26960
R;Sharples, K.; Plowman, G.D.; Rose, T.M.; Twardzik, D.R.; Purchio, A.F.
DNA 6, 239-244, 1987
A;Title: Cloning and sequence analysis of simian transforming growth factor-beta
cDNA.
A;Reference number: A26960; MUID:87246074; PMID:3474130
A;Accession: A26960
A;Molecule type: mRNA
A;Residues: 1-390 <SHA>
A;Cross-references: GB:M16658; NID:g176552; PIDN:AAA35369.1; PID:g176553
C;Superfamily: inhibin
C;Keywords: growth factor
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-390/Product: transforming growth factor beta #status predicted <MAT>

Query Match 89.6%; Score 1916.5; DB 2; Length 390;
Best Local Similarity 90.3%; Pred. No. 4.2e-146;
Matches 364; Conservative 10; Mismatches 16; Indels 13; Gaps 1;

```
Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTTPSRPAAGLSTCKTIDMELVKRKRIETIRGQILSKLRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | | | | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLSNRLLAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSKDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 6
JC4023
transforming growth factor beta-1 - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 24-Nov-1999
C;Accession: JC4023
R;Manning, A.M.; Auchampach, J.A.; Drong, R.F.; Slightom, J.L.
Gene 155, 307-308, 1995

A;Title: Cloning of a canine cDNA homologous to the human transforming growth factor-beta 1-encoding gene.
A;Reference number: JC4023; MUID:95237630; PMID:7721110
A;Accession: JC4023
A;Molecule type: mRNA
A;Residues: 1-390 <MAN>
A;Cross-references: GB:L34956; NID:g516071; PIDN:AAA51458.1; PID:g516072
C;Comment: This factor plays a multifunctional role as a regulator of mammalian cell growth and as a modulator of immune responses.
C;Genetics:
A;Gene: tgfbeta1
C;Superfamily: inhibin
C;Keywords: growth factor; transforming protein
F;288-390/Product: transforming growth factor beta 1 #status predicted <MAT>

Qy	1	MAPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Db	1	MPPSGRLRLPLLLPLLRLLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLS	60
Qy	61	SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPVPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVENTNKI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YEKVKKSPHSIYMLFNTSELREAVPEPVLLSRAELRLLRLKLKAEQHVELYQKYSNDSWR	180
Qy	181	YLSNRLLAPS DSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPS DTPEWLSFDVTGVVRQWL SHGGEVEGFRLSAHCSDSKDNTLQVDINGFS	240
Qy	241	SGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS	300
Db	241	SSRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRQRR-----ALDTNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

A;Title: The murine transforming growth factor-beta precursor.
A;Reference number: A01396; MUID:86168129; PMID:3007454
A;Accession: A01396
A;Molecule type: mRNA
A;Residues: 1-390 <DER>
A;Cross-references: GB:M13177; NID:g201952; PIDN:AAA40423.1; PID:g201953
A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide
C;Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide; the active molecule is a dimer of identical polypeptide chains linked by an interchain disulfide bond.
C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen; transformation
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-278/Domain: propeptide #status predicted <PRO>
F;244-246/Region: cell attachment (R-G-D) motif
F;279-390/Product: transforming growth factor beta-1 #status predicted <MAT>
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.9%; Score 1837.5; DB 1; Length 390;
Best Local Similarity 85.9%; Pred. No. 9.2e-140;
Matches 346; Conservative 15; Mismatches 29; Indels 13; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLPWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGGLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | : | | | : | | | | | : | | | | | | | | | | | | | | | | | |
Db    121 YEKTKDISHSIYMFFNTSDIREAVPEPPLLSRAELRLQRLKSSVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | : | : | | | | | | | | | | | : : | : | | | | | | :
Db    181 YLGNRLLTPTDTPWLSFDVTGVVRQWLNQGDGIQGFRRSAHCSDSKDNKLHVEINGIS 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLGTIHDNRPFLLLMATPLERAQHLHSSRHRR-----ALDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 8
S10219
transforming growth factor beta-1 precursor - rat
N;Alternate names: TGF type 2; TGF-beta

Db 241 PKRRGDLGTIHDMNRPFLLLMATPLERAQHLHSSRHRR-----ALDTNYCFS 287

QY 301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
 |||

Db 288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

QY 361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
 |||

Db 348 PGASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 9

A40057

transforming growth factor beta-1 precursor - bovine (fragment)

N;Alternate names: beta-TGF; cartilage-inducing factor-A; EGF-dependent TGF or dEGF; MGF-b; milk growth factor b; TGF-type II

C;Species: Bos primigenius taurus (cattle)

C;Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 16-Jul-1999

C;Accession: A40057; A42320; A05284; A24322; B61439

R;Van Obberghen-Schilling, E.; Kondaiah, P.; Ludwig, R.L.; Sporn, M.B.; Baker, C.C.

Mol. Endocrinol. 1, 693-698, 1987

A;Title: Complementary deoxyribonucleic acid cloning of bovine transforming growth factor-beta1.

A;Reference number: A40057; MUID:91042552; PMID:3153459

A;Accession: A40057

A;Molecule type: mRNA

A;Residues: 1-315 <VAN>

A;Cross-references: GB:M36271; NID:g163747; PIDN:AAA30778.1; PID:g163748

R;Ogawa, Y.; Schmidt, D.K.; Dasch, J.R.; Chang, R.J.; Glaser, C.B.

J. Biol. Chem. 267, 2325-2328, 1992

A;Title: Purification and characterization of transforming growth factor-beta2.3 and -beta1.2 heterodimers from bovine bone.

A;Reference number: A42320; MUID:92129307; PMID:1733936

A;Accession: A42320

A;Molecule type: protein

A;Residues: 204-209,'X',211-217 <OGA>

R;Roberts, A.B.; Anzano, M.A.; Meyers, C.A.; Wideman, J.; Blacher, R.; Pan, Y.C.E.; Stein, S.; Lehrman, S.R.; Smith, J.M.; Lamb, L.C.; Sporn, M.B.

Biochemistry 22, 5692-5698, 1983

A;Title: Purification and properties of a type beta transforming growth factor from bovine kidney.

A;Reference number: A05284; MUID:84104793; PMID:6607069

A;Accession: A05284

A;Molecule type: protein

A;Residues: 204-218 <ROB>

R;Seyedin, S.M.; Thompson, A.Y.; Bentz, H.; Rosen, D.M.; McPherson, J.M.; Conti, A.; Siegel, N.R.; Galluppi, G.R.; Piez, K.A.

J. Biol. Chem. 261, 5693-5695, 1986

A;Title: Cartilage-inducing factor-A. Apparent identity to transforming growth factor-beta.

A;Reference number: A24322; MUID:86195954; PMID:3754555

A;Accession: A24322

A;Molecule type: protein

A;Residues: 204-233 <SEY>

R;Jin, Y.; Cox, D.A.; Knecht, R.; Raschdorf, F.; Cerletti, N.

J. Protein Chem. 10, 565-575, 1991

A;Title: Separation, purification, and sequence identification of TGF-beta1 and TGF-beta2 from bovine milk.
A;Reference number: A61439; MUID:92189724; PMID:1799413
A;Accession: B61439
A;Molecule type: protein
A;Residues: 204-209,'X',211-217,'XX',220-232 <JIN>
C;Comment: This polypeptide is composed of two polypeptide chains cross-linked by disulfide bonds. It has been found in neoplastic and non-neoplastic tissues.
C;Comment: Type II TGF does not bind to the EGF receptor and lacks intrinsic mitogenic activity, but in soft agar, it reacts synergistically with either type I TFG or EGF, and induces cell proliferation. Cells grown in monolayer do not respond in a similar manner to these growth factors, but morphologically do acquire a transformed phenotype.
C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor; heterodimer
F;204-315/Product: transforming growth factor beta-1 #status experimental <MAT>
F;7,61,101/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.6%; Score 1574.5; DB 2; Length 315;
Best Local Similarity 89.9%; Pred. No. 9e-119;
Matches 295; Conservative 9; Mismatches 11; Indels 13; Gaps 1;

```

Qy      76 AVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF 135
      |:||||||||||||||| | ||||||||||||||||| ||:|||| | : ||:| |
Db      1 AILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMFF 60

Qy     136 NTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWRYLNSRLLAPSDSPEW 195
      |||||||||||||||||::||| |||||||||||||:||||| |||||||
Db      61 NTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWRYLNSRLLAPSDSPEW 120

Qy     196 LSFDTVGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFNSGRRGDLATIHGMNR 255
      ||||||||||||||||| ||||||||| | ||||||| |:||||:||||| |||||||
Db     121 LSFDTVGVVRQWLTRREEIEGFRLSAHCSCDSKDNTLQVDINGFSSGRRGDLATIHGMNR 180

Qy     256 PFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDF 315
      ||||||||||||||||| ||||||| ||||||||||||||||||| |||||||
Db     181 PFLLLMATPLERAQHLHSSRHR-----ALDTNYCFSSTEKNCCVRQLYIDF 227

Qy     316 RKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALE 375
      ||||||||||||||||||| ||||||||||||||||||| |||||||||||||||
Db     228 RKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALE 287

Qy     376 PLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      |||||||||||||||||||
Db     288 PLPIVYYVGRKPKVEQLSNMIVRSCKCS 315

```

RESULT 10

A41918
transforming growth factor beta-4 precursor - chicken (fragment)
N;Alternate names: TGF-beta 4
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: A41918; A34941; S03110
R;Burt, D.W.; Jakowlew, S.B.
Mol. Endocrinol. 6, 989-992, 1992

A;Title: Correction: a new interpretation of a chicken transforming growth factor-beta 4 complementary DNA.
A;Reference number: A41918; MUID:92357039; PMID:1353860
A;Accession: A41918
A;Molecule type: mRNA
A;Residues: 1-373 <BUR>
A;Cross-references: GB:M31160; GB:X08012; GB:S41706; NID:gl262437; PIDN:AAB05637.1; PID:gl262438
A;Note: sequence extracted from NCBI backbone (NCBIN:110186, NCBIIP:110187)
A;Note: this report corrects and reinterprets the sequence from reference A34941 R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B. Mol. Endocrinol. 2, 1186-1195, 1988
A;Title: Complementary deoxyribonucleic acid cloning of a messenger ribonucleic acid encoding transforming growth factor beta-4 from chicken embryo chondrocytes.
A;Reference number: A34941; MUID:89112198; PMID:2464131
A;Accession: A34941
A;Molecule type: mRNA
A;Residues: 'MDPMSIGPKSCGGSPWRPPGTAPWSIGSRRATASSSCSTSSRVRAEVGGRAL',122-209,'D',211-373 <JAK>
A;Cross-references: EMBL:X08012
A;Note: this sequence has been corrected in A41918
C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor
F;1/Domain: signal sequence (fragment) #status predicted <SIG>
F;223-225/Region: cell attachment (R-G-D) motif
F;260-373/Product: transforming growth factor beta-4 #status predicted <MAT>
F;54,109,153/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.0%; Score 1049; DB 2; Length 373;
Best Local Similarity 53.5%; Pred. No. 1.8e-76;
Matches 207; Conservative 52; Mismatches 100; Indels 28; Gaps 8;

```

Qy      30 LSTCKTIDMELVKRKRIEAI R G Q I L S K L R L A S P P S Q G D V P P G P L P E A V L A L Y N S T R D R V A 89
      ||||: :||: |:|||||:|||||||: || : || |||: | |||||:: :
Db      2 LSTCQRDL E A A K K R I E A V R G Q I L S K L R L T A P P P A S E T P P R P L P D D V R A L Y N S T Q E L L K 61

Qy     90 GES-VEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPV 148
      : : | |: :|||: |: | : : : : ||:: :|| | | |
Db     62 QRARLRPPDPGDEYWAKELRRIPMETTWGAMEHWQPQSHSIFVFVNSRARRG-GRPT 120

Qy    149 LLSRAELRLLRLK L K-----VEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGV 203
      || |||||: || | ||:|||| | | |||| | : : |||||
Db    121 LLHRAELRMLRQKAAADSAGTEQRLELYQGYGNASWRYLHGRSVRATADDEWLSFDVTDA 180

Qy    204 VRQWLTRREAIEGFRLSAHSSSD--SKDNTLHVEINGFNSGRRGDLATIHGMMNR--PFL 258
      | |||: | : |:|| | : : | || :|||: :| :| ||:
Db    181 VHQWLSGSELLGVFKLSVHCPCCEMGPGHAEEMRISIEGFEQ-ORGDMQSI AKKHRRVPYV 239

Qy    259 LLMATPLERAQHLHSSRHRRALDNTDYKDDDDKALDNTNYCF--SSTEKNCCVRQLYIDFR 316
      | || | ||| |||: | || |||: ||| : ||||| |||||
Db    240 LAMALPAERANELHSARRRR-----DLDTDYCFGPGTDEKNCCVRPLYIDFR 286

Qy    317 KDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEP 376
      ||| ||||| |||||: ||||| ||||: ||||| ||||| ||||| :|
Db    287 KDLQWKWIHEPKGYMANFCMGPCPYIWSADTQYTKVLALYNQHNP GASAAPCCVPQTLDLP 346

```

Qy 377 LPIVYYVGRKPKVEQLSNMIVRSCKCS 403
 |||:||||| :|||||||:|:|||||
 Db 347 LPIIYYVGRNVRVEQLSNMVVRACKCS 373

RESULT 11

A34939

transforming growth factor beta-3 precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 16-Jul-1999

C;Accession: A34939; S25850; S36125; S36124; I51181

R;Jakowlew, S.B.; Dillard, P.J.; Kondaiah, P.; Sporn, M.B.; Roberts, A.B.

Mol. Endocrinol. 2, 747-755, 1988

A;Title: Complementary deoxyribonucleic acid cloning of a novel transforming growth factor-beta messenger ribonucleic acid from chick embryo chondrocytes.

A;Reference number: A34939; MUID:89096966; PMID:3211158

A;Accession: A34939

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-412 <JAK>

A;Cross-references: GB:M31154; NID:g212758; PIDN:AAA49089.1; PID:g212759

R;Burt, D.W.; Paton, I.R.; Dey, B.R.

J. Mol. Endocrinol. 7, 175-183, 1991

A;Title: Comparative analysis of human and chicken transforming growth factor-beta-2 and -beta-3 promoters.

A;Reference number: S25850; MUID:92134496; PMID:1840616

A;Accession: S25850

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-117 <BUR>

A;Cross-references: EMBL:X58127; NID:g63815; PIDN:CAA41128.1; PID:g63816

A;Accession: S36125

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 119-172 <BU2>

A;Cross-references: EMBL:X60055; NID:g396688; PIDN:CAA42653.1; PID:g396689

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

A;Accession: S36124

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 173-322,'ELPT',327-412 <BU3>

A;Cross-references: EMBL:X60091

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

R;Jakowlew, S.B.; Lechleider, R.; Geiser, A.G.; Kim, S.J.; Santa-Coloma, T.A.; Cubert, J.; Sporn, M.B.; Roberts, A.B.

Mol. Endocrinol. 6, 1285-1298, 1992

A;Title: Identification and characterization of the chicken transforming growth factor-beta 3 promoter.

A;Reference number: I51181; MUID:93024487; PMID:1406706

A;Accession: I51181

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-117 <JA2>

A;Cross-references: GB:S46000; NID:g257172; PIDN:AAB23575.1; PID:g257173

C;Genetics:

A;Introns: 216/1; 252/1; 309/2; 360/3
A;Note: list of introns may be incomplete
C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-300/Domain: propeptide #status predicted <PRO>
F;301-412/Product: transforming growth factor beta-3 #status predicted <MAT>
F;74,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.8%; Score 873.5; DB 2; Length 412;
Best Local Similarity 46.6%; Pred. No. 2.5e-62;
Matches 198; Conservative 56; Mismatches 114; Indels 57; Gaps 14;

```

Qy      15 LLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGDVPPGGLP 74
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      9 LVLLSLLSFATVSLALSSCTTLDLEHIKKRVEAIRGQILSKLRLTSPPE--SVGPAHVP 66

Qy      75 EAVLALYNSTRDRVAGESVEPEPE-----PEADYYAKEVTRVLMVE---SGNQIYDKF 124
      : | | | | | | : | : | | | | | | | | | | | | | : | : | : |
Db      67 YQILALYNSTRELL--EEMEEKEEESCSQENTESEYYAKEIHKFDMIQGLPEHNELGICP 124

Qy     125 KGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND--- 177
      | | : : | | | | | | | | | | | : | : | : | : | |
Db     125 KGVTSNVFR-FNVS---SAEKNSTNLFRAEFRVLRVNPSSKRSEQRIELFQILRPDEHI 180

Qy     178 -SWRYLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGFRLSAH-----SSSDSKD 229
      | | | | | : | | | | | | | | | | | : | | | : | : |
Db     181 AKQRYLSGRNVQTRGSPEWLSFDVTDTVREWLLHRESNLGLEISIHCPCHTFQPNGDILE 240

Qy     230 N---TLHVEINGFNNG---RRGDLATI---HGMNRPFLLLMATPLERAQH--LHSSRHRR 278
      | | | | : : | : | | | | : : | : | | | | : | : | : |
Db     241 NLHEVLEIKFKGIDSEDDYGRGDLGRLKKQKDLHNPHLILMMLPPHRLESPTLGGQRKKR 300

Qy     279 ALDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGP 338
      | | | | | | : | : | | | | | | | | | | | | | | | | |
Db     301 -----ALDTNYCFRNLEENCCVRPLYIDFRQDLGWKQWVHEPKGYFANFCSGP 347

Qy     339 CPYIWSLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVR 398
      | | : | | : | | | | | | | | | | | | | | | | | : |
Db     348 CPYLRSAADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVK 407

Qy     399 SCKCS 403
      | | | |
Db     408 SCKCS 412

```

RESULT 12

B61036

transforming growth factor beta-5 precursor - African clawed frog

C;Species: *Xenopus laevis* (African clawed frog)

C;Date: 31-Dec-1993 #sequence_revision 03-Feb-1994 #text_change 16-Jul-1999

C;Accession: A34929; B61036

R;Kondaiah, P.; Sands, M.J.; Smith, J.M.; Fields, A.; Roberts, A.B.; Sporn, M.B.; Melton, D.A.

J. Biol. Chem. 265, 1089-1093, 1990

A;Title: Identification of a novel transforming growth factor-beta (TGF-beta5) mRNA in *Xenopus laevis*.

A;Reference number: A34929; MUID:90110090; PMID:2295601
A;Accession: A34929
A;Molecule type: mRNA
A;Residues: 1-382 <KON>
A;Cross-references: GB:J05180; NID:g214821; PIDN:AAA49968.1; PID:g214822
R;Roberts, A.B.; Rosa, F.; Roche, N.S.; Coligan, J.E.; Garfield, M.; Rebbert, M.L.; Kondaiah, P.; Danielpour, D.; Kehrl, J.H.; Wahl, S.M.; Dawid, I.B.; Sporn, M.B.
Growth Factors 2, 135-147, 1990
A;Title: Isolation and characterization of TGF-beta2 and TGF-beta5 from medium conditioned by Xenopus XTC cells.
A;Reference number: A61036; MUID:90253806; PMID:2340184
A;Accession: B61036
A;Molecule type: protein
A;Residues: 271-276, 'X', 278-284, 'XX', 287-299 <ROB>
C;Superfamily: inhibin
C;Keywords: growth factor
F;271-382/Product: transforming growth factor beta-5 #status experimental <MAT>

Query Match 40.7%; Score 870.5; DB 2; Length 382;
Best Local Similarity 46.3%; Pred. No. 4e-62;
Matches 190; Conservative 55; Mismatches 122; Indels 43; Gaps 11;

```

Qy      9 LPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RGQILSKLRLASPPSQGDV 68
      : : | ||| | |||| : ||| |::| ||||| ||||| : | |
Db      1 MEVLWMLLVLLVLHLSSLAMSLSTCKAVDMEEVRKRRIEAI RGQILSKLKLDKTPDV-DS 59

Qy     69 PPGPLPEAVLALYNSTRDRVAGESVEPE-----PEPEADYYAKEVTRVLMVESGNQIYDK 123
      : | : |||| : : : : | : ||||| : | | || : : |
Db     60 EKMTVPSEAIFLYNSTLEVIREKATREEEHVGHQDNIQDYAKQVYR--FESITELED- 115

Qy    124 FKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKL--KVEQHVELYQKYSNDSW-- 179
      | || | : || | || |||| : : : : : : | : || : | :
Db    116 -----HEFKFKFNASHVRENVGMNSLLHHAELRMYKKQTDKNMDQRMELFWKYQENGTT 170

Qy    180 -RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH--SSSDSKDNTLHVE 235
      ||| : : | ||:|||| | : || | | | | : : || : :
Db    171 SRYLESKYITPVTDDWMSFDVTKTVNEWLKRAEENEQFGLQPACKCPTPAKD----ID 226

Qy    236 INGFNSGRRGDLATI HGM--NRPFLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKAL 293
      | | ||||| : : : : : : | || : ||| : | :
Db    227 IEGF-PALRGDLASLSSKENTKPYLMITSMIPAERIDTVTSSRKKR-----GV 272

Qy    294 DTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVL 353
      ||| : ||||| : ||:||||| ||||| ||: || ||||| : |||||
Db    273 GQEYCFGNNGPNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCPYIWSMDTQYSKVL 332

Qy    354 ALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 403
      : ||||| : |||| | ||||| : |||| | ||||| : |||| |
Db    333 SLYNQNNPGASISPCCVPDVLEPLPIIYVGRGTAKVEQLSNMVVRSCNCS 382

```

RESULT 13
S01825
transforming growth factor beta-3 precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999

C;Accession: S01825
R;Derynck, R.; Lindquist, P.B.; Lee, A.; Wen, D.; Tamm, J.; Graycar, J.L.; Rhee, L.; Mason, A.J.; Miller, D.A.; Coffey, R.J.; Moses, H.L.; Chen, E.Y.
EMBO J. 7, 3737-3743, 1988
A;Title: A new type of transforming growth factor-beta, TGF-beta3.
A;Reference number: S01824; MUID:89091120; PMID:3208746
A;Accession: S01825
A;Molecule type: mRNA
A;Residues: 1-409 <DER>
A;Cross-references: EMBL:X14150; NID:g2127; PIDN:CAA32363.1; PID:g2128
C;Superfamily: inhibin
C;Keywords: growth factor
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-297/Domain: propeptide #status predicted <PRO>
F;298-409/Product: transforming growth factor beta-3 #status predicted <MAT>

Query Match 39.1%; Score 836; DB 2; Length 409;
Best Local Similarity 45.0%; Pred. No. 2.6e-59;
Matches 190; Conservative 58; Mismatches 122; Indels 52; Gaps 14;

```

Qy      15 LLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVPPGGLP 74
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      7  LVVLALLNFATVSLSMSTCTTLDLDFDHIKRKRVEAIRGQILSKLRLTSPDPDPSML--ANIP 64

Qy      75 EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMV--ESGNQIYDKFKG 126
      || ||||| : | || : : | : ||||| : : | : | : | : ||
Db      65 TQVLDLYNSTRELLVEVHGERGDDCTQENTESEYYAKEIYKFDMIQGLEEHNDLAVCPKG 124

Qy     127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND---S 178
      : : || | : : | ||| | : || : : | : || : || : |
Db     125 ITSKIFR-FNVSSVEK---NETNLFRAEFRVLMPNPSSKRSEQRIELFQILQPDEHIAK 180

Qy     179 WRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSSDSKDN- 230
      || : : | : ||||| || : || || : | : | : | : | : |
Db     181 QRYIDGKNLPTRGAAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENI 240

Qy     231 --TLHVEINGFNS---GRRGDLATIHGM--NRPFLLMATPLERAQH--LHSSRHRRALD 281
      : : : | : | |||| : : | | : || | : | : | : | : |
Db     241 QEVMEIKFKGVDSDDPGRGDLGRLKKKKEHSPHLILMMIIPDRLDNPLGAQRKKR--- 297

Qy     282 TNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPY 341
      ||||| : | : |||| | ||||| : ||||| : ||||| ||||
Db     298 -----ALDTNYCFRNLEENCCVRPLYIDFRQDLGWKWVHEPKGYANFCSGPCPY 347

Qy     342 IWSLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCK 401
      : | || : | || || || |||| : |||| | ||| : |||| | ||||| : ||||
Db     348 LRSADTTHSSVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTAKVEQLSNMIVKSK 407

Qy     402 CS 403
      ||
Db     408 CS 409

```

RESULT 14
A41397
transforming growth factor beta-3 precursor - mouse
C;Species: Mus musculus (house mouse)

C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 16-Jul-1999
 C;Accession: A41397; A61039; A61225
 R;Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.
 Mol. Endocrinol. 3, 1926-1934, 1989
 A;Title: Complementary DNA cloning of the murine transforming growth factor-
 beta3 (TGFbeta3) precursor and the comparative expression of TGFbeta3 and
 TGFbeta1 messenger RNA in murine embryos and adult tissues.
 A;Reference number: A41397; MUID:90190650; PMID:2628730
 A;Accession: A41397
 A;Molecule type: mRNA
 A;Residues: 1-410 <MIL>
 A;Cross-references: GB:M32745; NID:g201949; PIDN:AAA40422.1; PID:g201950
 R;Denhez, F.; Lafyatis, R.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.
 Growth Factors 3, 139-146, 1990
 A;Title: Cloning by polymerase chain reaction of a new mouse TGF-beta, mTGF-
 beta3.
 A;Reference number: A61039; MUID:91000714; PMID:2206556
 A;Accession: A61039
 A;Molecule type: mRNA
 A;Residues: 1-410 <DEN>
 R;Watrin, F.; Scotto, L.; Assoian, R.K.; Wolgemuth, D.J.
 Cell Growth Differ. 2, 77-83, 1991
 A;Title: Cell lineage specificity of expression of the murine transforming
 growth factor beta-3 and transforming growth factor beta-1 genes.
 A;Reference number: A61225; MUID:91299576; PMID:2069871
 A;Accession: A61225
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 285-410 <WAT>
 C;Superfamily: inhibin
 C;Keywords: glycoprotein; growth factor; growth regulation
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-298/Domain: propeptide #status predicted <PRO>
 F;259-261/Region: cell attachment (R-G-D) motif
 F;299-410/Product: transforming growth factor beta-3 #status predicted <MAT>
 F;72,133,140/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.9%; Score 832.5; DB 2; Length 410;
 Best Local Similarity 44.7%; Pred. No. 4.9e-59;
 Matches 188; Conservative 58; Mismatches 126; Indels 49; Gaps 13;

Qy	15	LLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLASPPSQGDVPPGPLP	74
		: :	:	:
Db	7	LVVLALLNLATISLSLSTCTTLD	FGHIKKKRVEAIRGQILSKLRLTSPPEPSVMT--HVP	64
Qy	75	EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMVE---	SGNQIYDKFKG	126
			: : : : : : :	
Db	65	YQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFDMIQGLAEHNELAVCPKG		124
Qy	127	TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---	KLKVEQHVELYQKYSND---S	178
		: : : : : :		
Db	125	ITSKVFR-FNVSSVEK---NGTNLFRAEFRVLRVPNPSSKRTEQRIELFQILRPDEHIAK		180
Qy	179	WRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSSDSKDN-		230
		: : : : :		
Db	181	QRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENV		240

```

Qy      231 --TLHVEINGFNS---GRRGDLATIHGM--NRPFLLLMATPLERAQHLHSSRHRRALDT 282
      :  ::  |  ::  |||  :  :  |  |  |  |  |  |  :
Db      241 HEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHNNPHLIILMMIPPHRLDSPGQGSQRK---- 296

Qy      283 NDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYI 342
      :  |||  |||  :  |  |||  |||  |||  |||  |||  |||  |||  |||  :
Db      297 -----KRALDTNYCFRNLEENCCVRPLYIDFRQDLGWKQVHEPKGYANFCSGPCPYL 349

Qy      343 WSLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKC 402
      |  |  :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      350 RSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCK 409

Qy      403 S 403
      |
Db      410 S 410

```

RESULT 15

A55706

transforming growth factor beta-3 precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 16-Jul-1999

C;Accession: A55706; B40699; S36042

R;Wang, J.; Kuliszewski, M.; Yee, W.; Sedlackova, L.; Xu, J.; Tseu, I.; Post, M.
J. Biol. Chem. 270, 2722-2728, 1995

A;Title: Cloning and expression of glucocorticoid-induced genes in fetal rat
lung fibroblasts. Transforming growth factor-beta-3.

A;Reference number: A55706; MUID:95155340; PMID:7852342

A;Accession: A55706

A;Molecule type: mRNA

A;Residues: 1-410 <WAN>

A;Cross-references: GB:U03491

A;Note: it is uncertain whether Met-1 is the initiator

R;McKinnon, R.D.; Piras, G.; Ida Jr., J.A.; Dubois-Dalcq, M.
J. Cell Biol. 121, 1397-1407, 1993

A;Title: A role for TGF-beta in oligodendrocyte differentiation.

A;Reference number: A40699; MUID:93286190; PMID:8509457

A;Accession: B40699

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 157-211 <MCK>

A;Cross-references: EMBL:X71903; NID:g311326; PIDN:CAA50722.1; PID:g311327

C;Superfamily: inhibin

```

Query Match          38.9%;  Score 831.5;  DB 2;  Length 410;
Best Local Similarity 44.7%;  Pred. No. 5.9e-59;
Matches 188;  Conservative 58;  Mismatches 126;  Indels 49;  Gaps 13;

```

```

Qy      15 LLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVPPGGLP 74
      |  :  |  :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      7  LVVLALLNLATVSLSLSTCTTLDGHIKKKRVEAIRGQILSKLRLTSPPEPSVMT--HVP 64

Qy      75 EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMVE---SGNQIYDKFKG 126
      |||  |||  :  :  ||  |  :  |  :  |||  ||  :  |  :  ||  ||
Db      65 YQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFDMIQGLAEHNELAVCPKG 124

Qy      127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLRL----KLKVEQHVELYQKYSND----S 178

```

```

      :: || | : : | ||| |:|: : || :||: |
Db      125 ITSKVFR-FNVSSVEK---NGTNLFRAEFRVL RVPNPSSKRTEQRIELFQILRPDEHIAK 180
Qy      179 WRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSSDSKDN- 230
      ||: : | : ||||| ||:| |||: | :| | : | :|
Db      181 QRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENV 240
Qy      231 --TLHVEINGFNS---GRRGDLATIHGM---NRPFLLMATPLERAQHLHSSRHRRALDT 282
      : :: | :: ||| : : | |:| | | | :
Db      241 HEVMEIKFKGVNEDDHGRGDLGRLKKQKDHHNPHLILMMIPPHRLDSPGQGGQRK---- 296
Qy      283 NDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYI 342
      : ||||| : |:| ||| |||||:| ||||:| ||||:| ||||:
Db      297 -----KRALDTNYCFRNLEENCCVRPLYIDFRQDLGWKWVHEPKGYANFCSGPCPYL 349
Qy      343 WSLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKC 402
      | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      350 RSSDTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCK 409
Qy      403 S 403
      |
Db      410 S 410

```

Search completed: October 28, 2003, 09:09:49
Job time : 14.3169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 00:08:41 ; Search time 8.07747 Seconds
(without alignments)
2346.251 Million cell updates/sec

Title: US-10-017-372E-33
Perfect score: 2139
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2032.5	95.0	390	1	TGF1_PIG	P07200 sus scrofa
2	1931.5	90.3	390	1	TGF1_SHEEP	P50414 ovis aries
3	1918.5	89.7	390	1	TGF1_HUMAN	P01137 homo sapien
4	1916.5	89.6	390	1	TGF1_CERAE	P09533 cercopithec
5	1895.5	88.6	390	1	TGF1_CANFA	P54831 canis famil
6	1841.5	86.1	390	1	TGF1_HORSE	O19011 equus cabal
7	1837.5	85.9	390	1	TGF1_MOUSE	P04202 mus musculu
8	1837.5	85.9	390	1	TGF1_RAT	P17246 rattus norv
9	1828.5	85.5	390	1	TGF1_CAVPO	Q9zly6 cavia porce
10	1574.5	73.6	315	1	TGF1_BOVIN	P18341 bos taurus
11	1049	49.0	373	1	TGF1_CHICK	P09531 gallus gall
12	870.5	40.7	382	1	TGF1_XENLA	P16176 xenopus lae
13	854.5	39.9	412	1	TGF3_CHICK	P16047 gallus gall
14	836	39.1	409	1	TGF3_PIG	P15203 sus scrofa
15	832.5	38.9	410	1	TGF3_MOUSE	P17125 mus musculu
16	831.5	38.9	412	1	TGF3_RAT	Q07258 rattus norv
17	829.5	38.8	412	1	TGF3_HUMAN	P10600 homo sapien
18	829	38.8	414	1	TGF2_MOUSE	P27090 mus musculu
19	814	38.1	414	1	TGF2_HUMAN	P08112 homo sapien
20	813.5	38.0	413	1	TGF2_XENLA	P17247 xenopus lae
21	813	38.0	412	1	TGF2_CHICK	P30371 gallus gall
22	810	37.9	435	1	TGF2_PIG	P09858 sus scrofa
23	810	37.9	442	1	TGF2_RAT	Q07257 rattus norv
24	482	22.5	112	1	TGF2_BOVIN	P21214 bos taurus
25	327.5	15.3	375	1	GDF8_MELGA	O42221 meleagris g
26	325.5	15.2	375	1	GDF8_CHICK	O42220 gallus gall
27	315.5	14.7	375	1	GDF8_PIG	O18831 sus scrofa
28	314.5	14.7	375	1	GDF8_PAPHA	O18828 papio hamad
29	312.5	14.6	376	1	GDF8_RAT	O35312 rattus norv
30	311.5	14.6	375	1	GDF8_HUMAN	O14793 homo sapien
31	310.5	14.5	376	1	GDF8_MOUSE	O08689 mus musculu
32	309.5	14.5	375	1	GDF8_SHEEP	O18830 ovis aries
33	304	14.2	405	1	GDFB_MOUSE	Q9z1w4 mus musculu
34	304	14.2	407	1	GDFB_HUMAN	O95390 homo sapien
35	300.5	14.0	375	1	GDF8_BOVIN	O18836 bos taurus
36	284.5	13.3	374	1	GDF8_BRARE	O42222 brachydanio
37	282.5	13.2	345	1	GDFB_RAT	Q9z217 rattus norv
38	279	13.0	425	1	IHBA_SHEEP	P43032 ovis aries
39	276.5	12.9	424	1	IHBA_PIG	P03970 sus scrofa
40	274.5	12.8	424	1	IHBA_MOUSE	Q04998 mus musculu
41	274.5	12.8	424	1	IHBA_RAT	P18331 rattus norv
42	274	12.8	425	1	IHBA_BOVIN	P07995 bos taurus
43	273.5	12.8	426	1	IHBA_HORSE	P55102 equus cabal
44	272.5	12.7	426	1	IHBA_HUMAN	P08476 homo sapien
45	269.5	12.6	424	1	IHBA_CHICK	P27092 gallus gall

ALIGNMENTS

RESULT 1

TGF1_PIG

ID TGF1_PIG STANDARD; PRT; 390 AA.
AC P07200; P08832;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1).
GN TGFB1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=87174844; PubMed=3470708;
RA Derynck R., Rhee L.;
RT "Sequence of the porcine transforming growth factor-beta precursor.";
RL Nucleic Acids Res. 15:3187-3187(1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT VAL-114.
RC STRAIN=Miniature swine;
RX MEDLINE=89054010; PubMed=2461367;
RA Kondaiah P., van Obberghen-Schilling E., Ludwig R.L., Dhar R.,
RA Sporn M.B., Roberts A.B.;
RT "cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.
RT Evidence for alternate splicing and polyadenylation.";
RL J. Biol. Chem. 263:18313-18317(1988).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT VAL-114.
RX MEDLINE=88335639; PubMed=3166520;
RA Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;
RT "Nucleotide sequence of chicken transforming growth factor-beta 1
RT (TGF-beta 1).";
RL Nucleic Acids Res. 16:8730-8730(1988).
RN [4]
RP SHOWS THAT REF.3 SEQUENCE IS FROM PIG.
RA Jakowlew S.B.;
RL Unpublished observations (MAR-1996).
RN [5]
RP SEQUENCE FROM N.A., AND VARIANT VAL-114.
RA Wimmers K., Chomdej S., Ponsuksili S., Schellander K.;
RT "Polymorphism in the porcine transforming growth factor beta 1
RT gene.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE OF 279-322.
RX MEDLINE=87102890; PubMed=2879635;
RA Cheifetz S., Weatherbee J.A., Tsang M.L.S., Anderson J.K., Mole J.E.,
RA Lucas R., Massague J.;
RT "The transforming growth factor-beta system, a complex pattern of
RT cross-reactive ligands and receptors.";
RL Cell 48:409-415(1987).
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL

CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -!- CAUTION: REF.3 SEQUENCE WHICH WAS SAID TO ORIGINATE FROM CHICKEN
 CC WHITE LEGHORN, SEEMS (REF.4) TO ORIGINATE FROM PIG.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Y00111; CAA68291.1; -.
 DR EMBL; M23703; AAA64616.1; -.
 DR EMBL; X12373; CAA30933.1; -.
 DR EMBL; AF461808; AAL57902.1; -.
 DR PIR; A27512; A27512.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal; Polymorphism.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 278
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
 FT VARIANT 114 114 L -> V.
 FT CONFLICT 6 7 LR -> PG (IN REF. 3).
 FT CONFLICT 180 180 R -> G (IN REF. 3).
 FT CONFLICT 237 237 N -> NA (IN REF. 3).
 SQ SEQUENCE 390 AA; 44294 MW; A6E2C3659FC384E6 CRC64;

Query Match 95.0%; Score 2032.5; DB 1; Length 390;
 Best Local Similarity 95.8%; Pred. No. 1.9e-152;

Matches 386; Conservative 1; Mismatches 3; Indels 13; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMLESQNI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRR-----ALDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 2

TGF1_SHEEP

```

ID   TGF1_SHEEP      STANDARD;          PRT;   390 AA.
AC   P50414;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transforming growth factor beta 1 precursor (TGF-beta 1).
GN   TGFB1.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95121932; PubMed=7821809;
RA   Woodall C.J., McLaren L.J., Watt N.J.;
RT   "Sequence and chromosomal localisation of the gene encoding ovine
RT   latent transforming growth factor-beta 1.";
RL   Gene 150:371-373(1994).
RN   [2]
RP   SEQUENCE OF 281-390 FROM N.A.
RC   STRAIN=Merino; TISSUE=Skin;
RX   MEDLINE=95268698; PubMed=7749621;
RA   Sutton R., Ward W.G., Raphael K.A., Cam G.R.;

```

```

RT  "Growth factor expression in skin during wool follicle development.";
RL  Comp. Biochem. Physiol. 110B:697-705(1995).
CC  -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC  -!- SUBUNIT: Homodimer; disulfide-linked.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the TGF-beta family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X76916; CAA54242.1; -.
DR  EMBL; L36038; AAA31526.1; -.
DR  PIR; I46463; I46463.
DR  HSSP; P01137; 1KLA.
DR  InterPro; IPR002400; GF_cysknot.
DR  InterPro; IPR003911; TGF_TGFb.
DR  InterPro; IPR001839; TGFb.
DR  InterPro; IPR001111; TGFb_N.
DR  Pfam; PF00019; TGF-beta; 1.
DR  Pfam; PF00688; TGFb_propeptide; 1.
DR  PRINTS; PR00438; GFCYSKNOT.
DR  PRINTS; PR01423; TGFBETA.
DR  ProDom; PD000357; TGFb; 1.
DR  SMART; SM00204; TGFB; 1.
DR  PROSITE; PS00250; TGF_BETA_1; 1.
KW  Growth factor; Mitogen; Glycoprotein; Signal.
FT  SIGNAL          1      23      POTENTIAL.
FT  PROPEP          24     278      POTENTIAL.
FT  CHAIN           279     390      TRANSFORMING GROWTH FACTOR BETA 1.
FT  DISULFID        285     294      BY SIMILARITY.
FT  DISULFID        293     356      BY SIMILARITY.
FT  DISULFID        322     387      BY SIMILARITY.
FT  DISULFID        326     389      BY SIMILARITY.
FT  DISULFID        355     355      INTERCHAIN (BY SIMILARITY).
FT  CARBOHYD        82      82      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD       136     136      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD       176     176      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  SITE           244     246      CELL ATTACHMENT SITE (POTENTIAL).
SQ  SEQUENCE       390 AA;  44291 MW;  1C247299484D0E57 CRC64;

```

```

Query Match          90.3%;  Score 1931.5;  DB 1;  Length 390;
Best Local Similarity 90.6%;  Pred. No. 1.7e-144;
Matches 365;  Conservative 10;  Mismatches 15;  Indels 13;  Gaps 1;

```

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKKRIEAI RQILSKLRLA 60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLMLTPGRPVAGLSTCKTIDMELVKKRGIEAI RQILSKLRLA 60

```



```

QY      61 SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 SPPSQGDVPPGGLPEAILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKI 120

QY     121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLKVEQHVELYQKYSNDSWR 180
      |||:|:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     121 YDKMKSSSHSIYMFFNTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWR 180

QY     181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTHREEIEGFRLSAHCSCDSKDNTLQVDINGFS 240

QY     241 SGRRGDLATIHGMNRPFLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     241 SGRRGDLATIHGMNRPFLLMATPLERAQHLHSSRHR-----ALDTNYCFS 287

QY     301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

QY     361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 3

TGF1_HUMAN

```

ID   TGF1_HUMAN      STANDARD;          PRT;   390 AA.  .
AC   P01137; Q9UCG4;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Transforming growth factor beta 1 precursor (TGF-beta 1).
GN   TGFBI OR TGFB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87174845; PubMed=3470709;
RA   Derynck R., Rhee L., Chen E.Y., van Tilburg A.;
RT   "Intron-exon structure of the human transforming growth factor-beta
RT   precursor gene.";
RL   Nucleic Acids Res. 15:3188-3189(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85296301; PubMed=3861940;
RA   Derynck R., Jarrett J.A., Chen E.Y., Eaton D.H., Bell J.R.,
RA   Assoian R.K., Roberts A.B., Sporn M.B., Goeddel D.V.;
RT   "Human transforming growth factor-beta complementary DNA sequence and
RT   expression in normal and transformed cells.";
RL   Nature 316:701-705(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Duodenum, and Eye;

```

RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 279-390 FROM N.A.
 RC TISSUE=Carcinoma;
 RA Urushizaki Y., Niitsu Y., Terui T., Koshida Y., Mahara K., Kohgo Y.,
 RA Urushizaki I., Takahashi Y., Ito H.;
 RT "Cloning and expression of the gene for human transforming growth
 RT factor-beta in Escherichia coli.";
 RL Tumor Res. 22:41-55(1987).
 RN [5]
 RP SEQUENCE OF 279-329.
 RC TISSUE=Bladder carcinoma;
 RX MEDLINE=93229900; PubMed=8471846;
 RA Bourdrel L., Lin C.-H., Lauren S.L., Elmore R.H., Sugarman B.J.,
 RA Hu S., Westcott K.R.;
 RT "Recombinant human transforming growth factor-beta 1: expression by
 RT Chinese hamster ovary cells, isolation, and characterization.";
 RL Protein Expr. Purif. 4:130-140(1993).
 RN [6]
 RP SEQUENCE OF 279-301.
 RX MEDLINE=85131019; PubMed=2982829;
 RA Massague J., Like B.;
 RT "Cellular receptors for type beta transforming growth factor. Ligand
 RT binding and affinity labeling in human and rodent cell lines.";
 RL J. Biol. Chem. 260:2636-2645(1985).
 RN [7]
 RP STRUCTURE BY NMR OF 279-390.
 RX MEDLINE=93144319; PubMed=8424942;
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,
 RA Torchia D.A.;
 RT "Transforming growth factor beta 1: NMR signal assignments of the
 RT recombinant protein expressed and isotopically enriched using Chinese
 RT hamster ovary cells.";
 RL Biochemistry 32:1152-1163(1993).
 RN [8]
 RP STRUCTURE BY NMR OF 279-390.

RX MEDLINE=93144320; PubMed=8424943;
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,
 RA Torchia D.A.;
 RT "Transforming growth factor beta 1: secondary structure as determined
 RT by heteronuclear magnetic resonance spectroscopy.";
 RL Biochemistry 32:1164-1171(1993).
 RN [9]
 RP STRUCTURE BY NMR OF 279-390.
 RX MEDLINE=96266150; PubMed=8679613;
 RA Hinck A.P., Archer S.J., Qian S.W., Roberts A.B., Sporn M.B.,
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,
 RA Torchia D.A.;
 RT "Transforming growth factor beta 1: three-dimensional structure in
 RT solution and comparison with the X-ray structure of transforming
 RT growth factor beta 2.";
 RL Biochemistry 35:8517-8534(1996).
 CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,
 CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
 CC CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE
 CC SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
 CC POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X05839; CAA29283.1; -.
 DR EMBL; X05840; CAA29283.1; JOINED.
 DR EMBL; X05843; CAA29283.1; JOINED.
 DR EMBL; X05844; CAA29283.1; JOINED.
 DR EMBL; X05849; CAA29283.1; JOINED.
 DR EMBL; X05850; CAA29283.1; JOINED.
 DR EMBL; X02812; CAA26580.1; ALT_SEQ.
 DR EMBL; BC001180; AAH01180.1; -.
 DR EMBL; BC000125; AAH00125.1; -.
 DR EMBL; BC022242; AAH22242.1; -.
 DR EMBL; M38449; AAA36735.1; -.
 DR PIR; A27513; WFHU2.
 DR PDB; 1KLA; 17-AUG-96.
 DR PDB; 1KLC; 17-AUG-96.
 DR PDB; 1KLD; 17-AUG-96.
 DR Genew; HGNC:11766; TGFB1.
 DR MIM; 190180; -.
 DR GO; GO:0006916; P:anti-apoptosis; TAS.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.

```

DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFEBETA.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal; 3D-structure.
FT      SIGNAL          1      23      POTENTIAL.
FT      PROPEP          24      278
FT      CHAIN           279      390      TRANSFORMING GROWTH FACTOR BETA 1.
FT      DISULFID        285      294
FT      DISULFID        293      356
FT      DISULFID        322      387
FT      DISULFID        326      389
FT      DISULFID        355      355      INTERCHAIN.
FT      CARBOHYD        82      82      N-LINKED (GLCNAC. . .) (POTENTIAL)
FT      CARBOHYD       136     136      N-LINKED (GLCNAC. . .) (POTENTIAL)
FT      CARBOHYD       176     176      N-LINKED (GLCNAC. . .) (POTENTIAL)
FT      SITE           244     246      CELL ATTACHMENT SITE (POTENTIAL).
FT      CONFLICT        10      10      L -> P (IN REF. 2).
FT      CONFLICT       159     159      R -> RR (IN REF. 2).
FT      STRAND          281     281
FT      TURN            282     287
FT      STRAND          294     296
FT      STRAND          300     300
FT      TURN            302     305
FT      STRAND          313     313
FT      STRAND          317     317
FT      STRAND          321     323
FT      HELIX           335     346
FT      TURN            348     349
FT      STRAND          358     370
FT      TURN            371     372
FT      STRAND          373     387
SQ      SEQUENCE       390 AA;  44341 MW;  75391614250288FE CRC64;

```

Query Match 89.7%; Score 1918.5; DB 1; Length 390;
Best Local Similarity 90.6%; Pred. No. 1.8e-143;
Matches 365; Conservative 10; Mismatches 15; Indels 13; Gaps 1;

[illegible]

```

Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      :|||||
Db      241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS 287

Qy      301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      |||||
Db      288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy      361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      |||||
Db      348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 4

TGF1_CERAE

```

ID   TGF1_CERAE          STANDARD;          PRT;   390 AA.
AC   P09533;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transforming growth factor beta 1 precursor (TGF-beta 1).
GN   TGFB1.
OS   Cercopithecus aethiops (Green monkey) (Grivet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=9534;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87246074; PubMed=3474130;
RA   Sharples K., Plowman G.D., Rose T.M., Twardzik D.R., Purchio A.F.;
RT   "Cloning and sequence analysis of simian transforming growth
RT   factor-beta cDNA.";
RL   DNA 6:239-244(1987).
CC   -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROL PROLIFERATION,
CC   DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
CC   CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM HAVE
CC   SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC   ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
CC   POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the TGF-beta family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; M16658; AAA35369.1; -.
DR   PIR; A26960; A26960.
DR   HSSP; P01137; 1KLA.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR003911; TGF_TGFB.

```

DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbBETA.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT PROPEP 17 278
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 390 AA; 44356 MW; DFF63E2BAB44320E CRC64;

Query Match 89.6%; Score 1916.5; DB 1; Length 390;
 Best Local Similarity 90.3%; Pred. No. 2.5e-143;
 Matches 364; Conservative 10; Mismatches 16; Indels 13; Gaps 1;

Qy 1 MAPSGRLRLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKKRKRIEAI RQILSKLRLA 60
 |
 Db 1 MPPSGRLRLPLLLPLWLLVLTTPSRPAAGLSTCKTIDMELVKKRKRIETIRGQILSKLRLA 60

 Qy 61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
 |||||:|||||
 Db 61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

 Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLKVEQHVELYQKYSNDSWR 180
 |||||: ||:|
 Db 121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLRLKLKVEQHVELYQKYSNNSWR 180

 Qy 181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
 |||||:|||||:|
 Db 181 YLSNRL LAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSDSKDNTLQVDINGFT 240

 Qy 241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
 :|||||
 Db 241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRR-----ALDTNYCFS 287

 Qy 301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
 |||||
 Db 288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

 Qy 361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
 |||||
 Db 348 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 5

TGF1_CANFA

ID TGF1_CANFA STANDARD; PRT; 390 AA.
AC P54831;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1).
GN TGFB1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Jugular vein endothelial;
RX MEDLINE=95237630; PubMed=7721110;
RA Manning A.M., Auchampach J.A., Drong R.F., Slightom J.L.;
RT "Cloning of a canine cDNA homologous to the human transforming growth
RT factor-beta 1-encoding gene.";
RL Gene 155:307-308(1995).
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the TGF-beta family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L34956; AAA51458.1; -.
DR PIR; JC4023; JC4023.
DR HSSP; P01137; 1KLA.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR003911; TGF_TGFB.
DR InterPro; IPR001839; TGFB.
DR InterPro; IPR001111; TGFB_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFB_propeptide; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRINTS; PR01423; TGFBETA.
DR ProDom; PD000357; TGFB; 1.
DR SMART; SM00204; TGFB; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Mitogen; Glycoprotein; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 278 BY SIMILARITY.
FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.

FT	DISULFID	285	294	BY SIMILARITY.
FT	DISULFID	293	356	BY SIMILARITY.
FT	DISULFID	322	387	BY SIMILARITY.
FT	DISULFID	326	389	BY SIMILARITY.
FT	DISULFID	355	355	INTERCHAIN.
FT	CARBOHYD	82	82	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	244	246	CELL ATTACHMENT SITE (POTENTIAL).
SQ	SEQUENCE	390 AA;	44185 MW;	EB4780E88B7B590E CRC64;

Query Match 88.6%; Score 1895.5; DB 1; Length 390;
Best Local Similarity 89.6%; Pred. No. 1.1e-141;
Matches 361; Conservative 12; Mismatches 17; Indels 13; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
		:	
Db	1	MPPSGRLRLPLLLPLLRLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLS	60
Qy	61	SPPSQGDVPVPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVL MVESGNQI	120
		: : :	
Db	61	SPPSQGEVPPVPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVL MVENTNKI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		: : :	
Db	121	YEKVKKSPHSIYMLFNTSELREAVPEPVLLSRAELRLLRLKLKA EQHVELYQKYSNDSWR	180
Qy	181	YLSNRLLAPS D S P E W L S F D V T G V V R Q W L T R R E A I E G F R L S A H S S D S K D N T L H V E I N G F N	240
		: : : : :	
Db	181	YLSNRLLAPS DT P E W L S F D V T G V V R Q W L S H G G E V E G F R L S A H C S C D S K D N T L Q V D I N G F S	240
Qy	241	SGRRGDLATI HG MN RP F L L L M AT P L E R A Q H L H S S R H R R A L D T N D Y K D D D D K A L D T N Y C F S	300
Db	241	SSRRGDLATI HG MN RP F L L L M AT P L E R A Q H L H S S R Q R R - - - - - - - - - - - - - - - - ALD T N Y C F S	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD T Q Y S K V L A L Y N Q H N	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD T Q Y S K V L A L Y N Q H N	347
Qy	361	PGASAAPCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS	390

```

RESULT 6
TGF1_HORSE
ID   TGF1_HORSE      STANDARD;          PRT;    390 AA.
AC   O19011;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transforming growth factor beta 1 precursor (TGF-beta 1).
GN   TGFb1.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI TaxID=9796;

```


RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph node;
 RX MEDLINE=98185507; PubMed=9524819;
 RA Penha-Goncalves M.N., Onions D.E., Nicolson L.;
 RT "Cloning and sequencing of equine transforming growth factor-beta 1
 RT (TGF beta-1) cDNA.";
 RL DNA Seq. 7:375-378(1997).
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X99438; CAA67801.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 278 BY SIMILARITY.
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 390 AA; 43974 MW; A86D715F44549691 CRC64;

Query Match 86.1%; Score 1841.5; DB 1; Length 390;
 Best Local Similarity 87.3%; Pred. No. 2e-137;
 Matches 352; Conservative 12; Mismatches 26; Indels 13; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		: :	
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRAQVAGESAETEPEPEADYYAKEVTRVLMVEKENEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:	
Db	121	YKTVETGSHSIYMFNTSELRAAVDPMLLSRAELRLLRLKLSVEQHVELYQKYSNNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGFN	240
		: : : :	
Db	181	YLSNRLLTPSDSPEWLSFDVTGVVRQWLSQGGAMEGFRLSAHCSCDSKDNTLRVGINGFS	240
Qy	241	SGRRGDLATI HGMNRPFLLL MATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS	300
Db	241	SSRRGDLATIDGMNRPFLLL MATPLERAQQLHSSRHRR-----ALDTNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPQVLEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 7

TGF1 MOUSE

ID TGF1 MOUSE STANDARD; PRT; 390 AA.

AC P04202:

DT 20-MAR-1987 (Rel. 04, Created)

DT 20-MAR-1987 (Rel. 04, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transforming growth factor beta 1 precursor (TGF-beta 1).

GN TGFB1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86168129; PubMed=3007454;

RA Derynck R., Jarrett J.A., Chen E.Y., Goeddel D.V.;

RT "The murine transforming growth factor-beta precursor.";

RL J. Biol. Chem. 261:4377-4379 (1986).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c;

RX MEDLINE=96096545; PubMed=8522200;

RA Guron C., Sudarshan C., Raghow R.;

RT "Molecular organization of the gene encoding murine transforming

```
RT    growth factor beta 1.";
```

RL Gene 165:325-326 (1995).

RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6, and NOD/LT; TISSUE=Spleen;
 RA Poirot L., Benoist C., Mathis D.;
 RT "Transforming growth factor-beta 1 sequence and expression: no
 RT difference between NOD/Lt and C57Bl/6 mouse strains.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M13177; AAA40423.1; -.
 DR EMBL; L42462; AAB00138.1; -.
 DR EMBL; L42456; AAB00138.1; JOINED.
 DR EMBL; L42457; AAB00138.1; JOINED.
 DR EMBL; L42458; AAB00138.1; JOINED.
 DR EMBL; L42459; AAB00138.1; JOINED.
 DR EMBL; L42460; AAB00138.1; JOINED.
 DR EMBL; L42461; AAB00138.1; JOINED.
 DR EMBL; AJ009862; CAA08900.1; -.
 DR PIR; A01396; WFMS2.
 DR HSSP; P01137; 1KLA.
 DR MGD; MGI:98725; Tgfb1.
 DR GO; GO:0005578; C:extracellular matrix; IDA.
 DR GO; GO:0006954; P:inflammatory response; IMP.
 DR GO; GO:0007515; P:lymph gland development; IMP.
 DR GO; GO:0008220; P:necrosis; IMP.
 DR GO; GO:0016202; P:regulation of myogenesis; IDA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 278
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.

FT	DISULFID	293	356	BY SIMILARITY.
FT	DISULFID	322	387	BY SIMILARITY.
FT	DISULFID	326	389	BY SIMILARITY.
FT	DISULFID	355	355	INTERCHAIN (BY SIMILARITY).
FT	CARBOHYD	82	82	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	244	246	CELL ATTACHMENT SITE (POTENTIAL).
SQ	SEQUENCE	390 AA; 44310 MW; 4381A51B711D689E CRC64;		

Query Match 85.9%; Score 1837.5; DB 1; Length 390;
 Best Local Similarity 85.9%; Pred. No. 4.1e-137;
 Matches 346; Conservative 15; Mismatches 29; Indels 13; Gaps 1;

Qy	1	MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLPWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVKVEQHVELYQKYSNDSWR	180
Db	121	YEKTKDISHSIYMFNTSDIREAVPEPPLLSRAELRLQLRKSSVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLGNRLLTPTDTPEWLSFDVTGVVRQWLNQDGIQGFRFSAHCSCDSKDNKLHVEINGIS	240
Qy	241	SGRRGDLATIHGMNRPFLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS	300
Db	241	PKRRGDLGTIHDNRPFLLMATPLERAQHLHSSRHRR-----ALDTNYCFS	287
Qy	301	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 8

TGF1_RAT

ID	TGF1_RAT	STANDARD;	PRT;	390 AA.
AC	P17246;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Transforming growth factor beta 1 precursor (TGF-beta 1).			
GN	TGFB1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			

RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Heart;
 RX MEDLINE=90272425; PubMed=2349108;
 RA Qian S.W., Kondaiah P., Roberts A.B., Sporn M.B.;
 RT "cDNA cloning by PCR of rat transforming growth factor beta-1.";
 RL Nucleic Acids Res. 18:3059-3059(1990).
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X52498; CAA36741.1; -.
 DR PIR; S10219; S10219.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 278
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 390 AA; 44329 MW; 5E21108ED50D853C CRC64;

Query Match 85.9%; Score 1837.5; DB 1; Length 390;
 Best Local Similarity 85.9%; Pred. No. 4.1e-137;
 Matches 346; Conservative 14; Mismatches 30; Indels 13; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLPWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | : | |
Db     61 SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | | : | | : | | | | | | | | | | | | | | | | | | | : | |
Db    121 YDKTKDITHSIYMFNTSDIREAVPEPPLLSRAELRLQRFKSTVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | : | | : | | | | | | | | | | | | | | | | | | | :
Db    181 YLGNRL LTPTDTPEWLSFDVTGVVRQWLNQGDGIQGFRLSAHCSDSKDNVLHVEINGIS 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLGTIHDNRPFLLLMATPLERAQHLHSSRHRR-----ALDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 9

TGF1_CAVPO

```

ID   TGF1_CAVPO      STANDARD;          PRT;   390 AA.
AC   Q9Z1Y6; Q9QZB3; Q9R148;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transforming growth factor beta 1 precursor (TGF-beta 1).
GN   TGFB1.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Hartley;
RA   Jeevan A., McMurray D.N., Yoshimura T.;
RT   "Guinea pig transforming growth factor-beta in peritoneal exudates
RT   after BCG vaccination.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 265-382 FROM N.A.
RX   MEDLINE=99144670; PubMed=10025978;
RA   Scarozza A.M., Ramsingh A.I., Wicher V., Wicher K.;
RT   "Spontaneous cytokine gene expression in normal guinea pig blood and
RT   tissues.";
RL   Cytokine 10:851-859(1998).
RN   [3]

```

RP SEQUENCE OF 279-371 FROM N.A.
 RC STRAIN=Hartley; TISSUE=Trachea;
 RA Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,
 RA Sekizawa K.;
 RT "Guinea-pig transforming growth factor-beta expression in injured
 RT tracheal epithelium.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,
 CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
 CC CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE
 CC SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
 CC POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF191297; AAF02780.1; -.
 DR EMBL; AF097509; AAC83807.1; -.
 DR EMBL; AF169347; AAD49347.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT PROPEP 25 278 POTENTIAL.
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
 FT CONFLICT 279 279 G -> P (IN REF. 3).
 FT CONFLICT 286 286 F -> S (IN REF. 2).
 FT CONFLICT 309 309 K -> E (IN REF. 2).
 FT CONFLICT 322 322 C -> R (IN REF. 2).

FT CONFLICT 350 350 A -> G (IN REF. 2).
SQ SEQUENCE 390 AA; 44328 MW; 1539F849BA0C0FF1 CRC64;

Query Match 85.5%; Score 1828.5; DB 1; Length 390;
Best Local Similarity 85.6%; Pred. No. 2.1e-136;
Matches 345; Conservative 15; Mismatches 30; Indels 13; Gaps 1;

```
Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAI RQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSRLRLLPLLLPLLWLLVLAPGRPASGLSTCKTIDMELVKRKRIEAI RQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEPEPDYYAKEVTRVLMVDNSHNI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YKSIETVAHSIYMFNTSELREAVDPDLLSRAELRMQRLKLNVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLSNQLLTPSDTPEWLSFDVTGVVRQWLSQGELEGFRFSAHCSCDSKDNTLRVEINGIG 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLAAIHGMNRPFLLLMATPLERAQHLHSSRHRR-----GLDTNYCFS 287

Qy    301 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    288 STEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 347

Qy    361 PGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    348 PGASAAPCCVPQALEPLPIVYVGRKAKVEQLSNMIVRSCKCS 390
```

RESULT 10

TGF1_BOVIN

ID TGF1_BOVIN STANDARD; PRT; 315 AA.
AC P18341;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1) (Fragment).
GN TGFβ1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91042552; PubMed=3153459;
RA van Obberghen-Schilling E., Kondaiah P., Ludwig R.L., Sporn M.B.,
RA Baker C.C.;
RT "Complementary deoxyribonucleic acid cloning of bovine transforming
RT growth factor-beta 1.";

RL Mol. Endocrinol. 1:693-698(1987).
 RN [2]
 RP SUBUNITS.
 RC TISSUE=Bone;
 RX MEDLINE=92129307; PubMed=1733936;
 RA Ogawa Y., Schmidt D.K., Dasch J.R., Chang R.J., Glaser C.B.;
 RT "Purification and characterization of transforming growth factor-beta
 RT 2.3 and -beta 1.2 heterodimers from bovine bone.";
 RL J. Biol. Chem. 267:2325-2328(1992).
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked. Heterodimers of TGF-beta 1/2
 CC have been found in bone.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M36271; AAA30778.1; -.
 DR PIR; A40057; A40057.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFb.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein.
 FT NON_TER 1 1
 FT PROPEP <1 203
 FT CHAIN 204 315 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 210 219 BY SIMILARITY.
 FT DISULFID 218 281 BY SIMILARITY.
 FT DISULFID 247 312 BY SIMILARITY.
 FT DISULFID 251 314 BY SIMILARITY.
 FT DISULFID 280 280 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 169 171 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 315 AA; 36269 MW; C2717A23D994E00E CRC64;

Query Match 73.6%; Score 1574.5; DB 1; Length 315;
 Best Local Similarity 89.9%; Pred. No. 1.5e-116;
 Matches 295; Conservative 9; Mismatches 11; Indels 13; Gaps 1;

```

Qy      76 AVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF 135
      |:||||||||||||||| | ||||||||||||||||| ||:|||| | : ||:|| |
Db      1 AILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMFF 60

Qy     136 NTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWRYLNSRLLAPSDSPEW 195
      |||||||||||||||||:|||||||||||||||:|||||||||||||||
Db     61 NTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWRYLNSRLLAPSDSPEW 120

Qy     196 LSFDTVGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFNSGRRGDLATIHGMNR 255
      ||||||||||||||||| ||||||||| | ||||||||| |:|||||:|||||||||||||||
Db     121 LSFDTVGVVRQWLTRREEIEGFRLSAHSCSDSKDNTLQVDINGFSSGRRGDLATIHGMNR 180

Qy     256 PFLLLMATPLERAQHLHSSRHRRALDNTDYKDDDDKALDNTNYCFSSTEKNCCVRQLYIDF 315
      ||||||||||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db     181 PFLLLMATPLERAQHLHSSRHR-----ALDNTNYCFSSTEKNCCVRQLYIDF 227

Qy     316 RKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQALE 375
      ||||||||||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db     228 RKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQALE 287

Qy     376 PLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      ||||||||||||||||| |||||||||
Db     288 PLPIVYYVGRKPKVEQLSNMIVRSCKCS 315
  
```

RESULT 11

TGF1_CHICK

```

ID  TGF1_CHICK      STANDARD;      PRT;      373 AA.
AC  P09531;
DT  01-MAR-1989 (Rel. 10, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 4)
DE  (Fragment).
GN  TGFB1.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=White leghorn;
RX  MEDLINE=89112198; PubMed=2464131;
RA  Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;
RT  "Complementary deoxyribonucleic acid cloning of a messenger
RT  ribonucleic acid encoding transforming growth factor beta 4 from
RT  chicken embryo chondrocytes.";
RL  Mol. Endocrinol. 2:1186-1195(1988).
RN  [2]
RP  REVISIONS.
RX  MEDLINE=92357039; PubMed=1353860;
RA  Burt D.W., Jakowlew S.B.;
  
```

```

RT  "Correction: a new interpretation of a chicken transforming growth
RT  factor-beta 4 complementary DNA.";
RL  Mol. Endocrinol. 6:989-992(1992).
CC  -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC  -!- SUBUNIT: Homodimer; disulfide-linked.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the TGF-beta family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M31160; AAB05637.1; -.
DR  PIR; A41918; A41918.
DR  HSSP; P01137; 1KLA.
DR  InterPro; IPR003911; TGF_TGFB.
DR  InterPro; IPR001839; TGFB.
DR  InterPro; IPR001111; TGFB_N.
DR  Pfam; PF00019; TGF-beta; 1.
DR  Pfam; PF00688; TGFB_propeptide; 1.
DR  PRINTS; PR01423; TGF_BETA.
DR  ProDom; PD000357; TGFB; 1.
DR  SMART; SM00204; TGFB; 1.
DR  PROSITE; PS00250; TGF_BETA_1; 1.
KW  Growth factor; Mitogen; Glycoprotein; Signal.
FT  NON_TER      1      1
FT  SIGNAL        <1      1      POTENTIAL.
FT  PROPEP        2      259     POTENTIAL.
FT  CHAIN         260     373     TRANSFORMING GROWTH FACTOR BETA 1.
FT  DISULFID      266     277     BY SIMILARITY.
FT  DISULFID      276     339     BY SIMILARITY.
FT  DISULFID      305     370     BY SIMILARITY.
FT  DISULFID      309     372     BY SIMILARITY.
FT  DISULFID      338     338     INTERCHAIN (BY SIMILARITY).
FT  CARBOHYD      54      54      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD     109     109     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD     153     153     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  SITE          224     226     CELL ATTACHMENT SITE (POTENTIAL).
SQ  SEQUENCE      373 AA;  42634 MW;  9903F3479C8552E5 CRC64;

```

```

Query Match          49.0%;  Score 1049;  DB 1;  Length 373;
Best Local Similarity 53.5%;  Pred. No. 3.9e-75;
Matches 207;  Conservative 52;  Mismatches 100;  Indels 28;  Gaps 8;

```

```

Qy      30 LSTCKTIDMELVKKRKRIEAIKQILSKLRLASPPSQGDVPPGGLPEAVLALYNSTRDRA 89
        ||||: :||:| |:|||||:|||||||:|:| :|| |||: | |||||:::
Db      2 LSTCQRLDLEAAKKKRIEAVRGQILSKLRLTAPPPASETPPRPLPDDVRALYNSTQELK 61

```

```

Qy      90 GES-VEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPV 148
      : : | | : : | | | : | : | : : : | | : : | | | |
Db      62 QRARLRPPPDGPDEYWAKELRRIPMETTWGAMEHWQPQSHSIFVFVNVSRARRG-GRPT 120

Qy      149 LLSRAELRLRLRLKLGK-----VEQHVELYQKYSNDSWRYLSNRL LAPSDSPEWLSFDVTGV 203
      | | | | | : | | | | : | | | | | | | | | | : : | | | | | |
Db      121 LLHRAELRMLRQKAAADSAGTEQRLELYQGYGNASWRYLHGRSVRATADDEWLSFDVTDA 180

Qy      204 VRQWLTRREAIEGFRLSAHSSSD---SKDNTLHVEINGFNSGRRGDLATIHGMNR--PFL 258
      | | | : | : | : | : : : : | | | : | | : : | : | : | :
Db      181 VHQWLSGSELLGVFKLSVHCPCEMGPGHAEEMRISIEGFQ-QRGDMQSI AKKHRRVPYV 239

Qy      259 LLMATPLERAQHLHSSRRHRRALDTNDYKDDDDKALDTNYCF--SSTEKNCCVRQLYIDFR 316
      | | | | | | | | | | | : | | | | | | | | | | | | | | | |
Db      240 LAMALPAERANELHSARRRR-----DLDTDYCFGPGTDEKNCCVRPLYIDFR 286

Qy      317 KDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEP 376
      | | | | | | | | | | | : | | | | | | | | | | | | | | | |
Db      287 KDLQWKWIHEPKGYMANFCMGPCPYIWSADTQYTKVLALYNQHNP GASAAPCCVPQTLDP 346

Qy      377 LPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      | | | : | | | | : | | | | | : | : | | |
Db      347 LPIIYYVGRNVRVEQLSNMVVRACKCS 373

```

RESULT 12

TGF1_XENLA

ID TGF1_XENLA STANDARD; PRT; 382 AA.

AC P16176;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 5).

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90110090; PubMed=2295601;

RA Kondaiah P., Sands M.J., Smith J.M., Fields A., Roberts A.B.,

RA Sporn M.B., Melton D.A.;

RT "Identification of a novel transforming growth factor-beta (TGF-beta

RT 5) mRNA in Xenopus laevis.";

RL J. Biol. Chem. 265:1089-1093(1990).

RN [2]

RP SEQUENCE FROM N.A.

RA Vempati U.D., Kondaiah P.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: IMPORTANT ROLE IN CERTAIN ASPECTS OF DIFFERENTIATION.

CC -!- SUBUNIT: Homodimer; disulfide-linked.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the TGF-beta family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; J05180; AAA49968.1; -.
DR EMBL; AF009335; AAB64441.1; -.
DR EMBL; AF009331; AAB64441.1; JOINED.
DR EMBL; AF009332; AAB64441.1; JOINED.
DR EMBL; AF009333; AAB64441.1; JOINED.
DR EMBL; AF009334; AAB64441.1; JOINED.
DR PIR; A34929; B61036.
DR HSSP; P01137; 1KLA.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR003911; TGF_TGFb.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRINTS; PR01423; TGFbBETA.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Mitogen; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 270
FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA 1.
FT DISULFID 277 286 BY SIMILARITY.
FT DISULFID 285 348 BY SIMILARITY.
FT DISULFID 314 379 BY SIMILARITY.
FT DISULFID 318 381 BY SIMILARITY.
FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 234 236 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 382 AA; 44200 MW; 1034621C917AAE15 CRC64;

Query Match 40.7%; Score 870.5; DB 1; Length 382;
Best Local Similarity 46.3%; Pred. No. 4.3e-61;
Matches 190; Conservative 55; Mismatches 122; Indels 43; Gaps 11;

QY 9 LPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDV 68
: :| || |||| | |||| :||| :::||| ||||| :| |
Db 1 MEVLWMLLVLLVLHLSSLAMSLSTCKAVDMEEVRKRRIEAIHQILSKLKLDKTPDV-DS 59

QY 69 PPGPLPEAVLALYNSTRDRVAGESVEPE-----PEPEADYYAKEVTRVLMVESGNQIYDK 123
: | : |||| : : : : | : ||||:| | || : :|
Db 60 EKMTVPSEAIFLYNSTLEVIREKATREEEHVGHQNIQDYAKQVYR---FESITELED- 115

QY 124 FKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKL--KVEQHVELYQKYSNDSW-- 179
| || | :|| | || ||||: : : : :| :||: || :
Db 116 -----HEFKFKFNASHVRENVGMNSLLHHAELRMYKKQTDKNMDQRMELFWKYQENGTH 170

QY 180 -RYLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH---SSSDSKDNTLHVE 235

```

      ||| :: : |   ||:|||| | :|| | | | | : :|| ::
Db      171 SRYLESKYITPVTDDDEWMSFDVTKTVNEWLKRAEENEQFGLQPACKCPTPQAKD----ID 226

Qy      236 INGFNSGRRGDLATIHGM--NRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKAL 293
      | ||   |||||::   :|:|:: : | ||   : ||| :|
Db      227 IEGF-PALRGDLASLSSKENTKPYLMITSMPAERIDTVTSSRKKR-----GV 272

Qy      294 DTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVL 353
      ||| :   ||||: |||:||||||||||||||||| ||:||| |||||:|||||||
Db      273 GQEYCFGNGPNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCPYIWSMDTQYSKVL 332

Qy      354 ALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      333 SLYNQNNPGASISPCCVPDVLEPLPIIYYVGRTAKVEQLSNMVVRSCNCS 382

```

RESULT 13

TGF3_CHICK

```

ID      TGF3_CHICK      STANDARD;      PRT;      412 AA.
AC      P16047;
DT      01-APR-1990 (Rel. 14, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 3 precursor (TGF-beta 3).
GN      TGFB3.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89096966; PubMed=3211158;
RA      Jakowlew S.B., Dillard P.J., Kondaiah P., Sporn M.B., Roberts A.B.;
RT      "Complementary deoxyribonucleic acid cloning of a novel transforming
RT      growth factor-beta messenger ribonucleic acid from chick embryo
RT      chondrocytes.";
RL      Mol. Endocrinol. 2:747-755(1988).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=White leghorn;
RX      MEDLINE=95169270; PubMed=7865129;
RA      Burt D.W., Dey B.R., Paton I.R., Morrice D.R., Law A.S.;
RT      "The chicken transforming growth factor-beta 3 gene: genomic
RT      structure, transcriptional analysis, and chromosomal location.";
RL      DNA Cell Biol. 14:111-123(1995).
RN      [3]
RP      SEQUENCE OF 1-117 FROM N.A.
RC      STRAIN=White leghorn; TISSUE=Blood;
RX      MEDLINE=92134496; PubMed=1840616;
RA      Burt D.W., Dey B.R., Paton I.R.;
RT      "Comparative analysis of human and chicken transforming growth
RT      factor-beta 2 and -beta 3 promoters.";
RL      J. Mol. Endocrinol. 7:175-183(1991).
RN      [4]
RP      SEQUENCE OF 1-117 FROM N.A.
RX      MEDLINE=93024487; PubMed=1406706;

```

RA Jakowlew S.B., Lechleider R., Geiser A.G., Kim S.J.,
 RA Santa-Coloma T.A., Cubert J., Sporn M.B., Roberts A.B.;
 RT "Identification and characterization of the chicken transforming
 RT growth factor-beta 3 promoter.";
 RL Mol. Endocrinol. 6:1285-1298(1992).
 CC -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M31154; AAA49089.1; -.
 DR EMBL; X58127; CAA41128.2; -.
 DR EMBL; X60055; CAA41128.2; JOINED.
 DR EMBL; X60091; CAA41128.2; JOINED.
 DR EMBL; X60090; CAA41128.2; JOINED.
 DR EMBL; S46000; AAB23575.1; -.
 DR PIR; A34939; A34939.
 DR HSSP; P10600; 1TGJ.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 300
 FT CHAIN 301 412 TRANSFORMING GROWTH FACTOR BETA 3.
 FT DISULFID 307 316 BY SIMILARITY.
 FT DISULFID 315 378 BY SIMILARITY.
 FT DISULFID 344 409 BY SIMILARITY.
 FT DISULFID 348 411 BY SIMILARITY.
 FT DISULFID 377 377 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 261 263 CELL ATTACHMENT SITE (POTENTIAL).
 FT CONFLICT 323 326 ELPT -> DFRQ (IN REF. 1).
 SQ SEQUENCE 412 AA; 47077 MW; 1CAB883170069D55 CRC64;

Query Match 39.9%; Score 854.5; DB 1; Length 412;
 Best Local Similarity 45.9%; Pred. No. 8.7e-60;
 Matches 195; Conservative 56; Mismatches 117; Indels 57; Gaps 14;

Qy	15	LLWLLVLTPGRPAAGLSTCKTIDMELVKKRIEAIRGQILSKLRLASPPSQGDVPPGPLP	74
Db	9	LVLLSLLSFATVSLALSSCTTLDLEHIKKKRVEAIRGQILSKLRLTSPPE--SVGPAHVP	66
Qy	75	EAVLALYNSTRDRVAGESVEPEPE-----PEADYYAKEVTRVLMVE---SGNQIYDKF	124
Db	67	YQILALYNSTRELL--EEMEEEKEESCSQENTESEYYAKEIHKFDMIQGLPEHNELGICP	124
Qy	125	KGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND---	177
Db	125	KGVTSNVFR-FNVS--SAEKNSTNLFRAEFRVLRVNPNSSKRSEQRIELFQILRPDEHI	180
Qy	178	-SWRYLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGFRLSAH-----SSSDSKD	229
Db	181	AKQRYLSGRNVQTRGSPWLSFDVTDTVREWLLHRESNLGLEISIHCPCHTFQPNGDILE	240
Qy	230	N---TLHVEINGFNNSG---RRGDLATI---HGMNRPFLLLMATPLERAQH--LHSSRHR	278
Db	241	NLHEVLEIKFKGIDSEDDYGRGDLGRLLKKQKDLHNPHLILMMLPPHRLESPTLGGQRKKR	300
Qy	279	ALDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGP	338
Db	301	-----ALDTNYCFRNLEENCCVRPLYIELPTDLGWKWVHEPKGYFANFCSGP	347
Qy	339	CPYIWSLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVR	398
Db	348	CPYLRSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYVVGRTPKVEQLSNMVVK	407
Qy	399	SCKCS 403	
Db	408	SCKCS 412	

RESULT 14

```

TGF3_PIG
ID   TGF3_PIG          STANDARD;          PRT;   409 AA.
AC   P15203;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transforming growth factor beta 3 precursor (TGF-beta 3).
GN   TGFB3.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=89091120; PubMed=3208746;
RA   Derynck R., Lindquist B., Lee A., Wen D., Tamm J., Graycar J.L.,
RA   Rhee L., Mason A.J., Miller D.A., Coffey R.J., Moses H.L.,
RA   Chen E.Y.;
RT   "A new type of transforming growth factor-beta, TGF-beta 3.";
RL   EMBO J. 7:3737-3743(1988).
CC   -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.

```



```

      ||:  : |      : ||||| ||: || ||:  |  :| |      : |  :|
Db      181 QRYIDGKNLPTRGAAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENI 240

Qy      231 --TLHVEINGFNS---GRRGDLATIHGM--NRPFLLLMATPLERAQH--LHSSRHRRALD 281
      : :: | :|      |||| :      : | |:|| | :| : | : | :|
Db      241 QEVMEIKFKGVDSEDDPGRGDLGRLKKKKEHSPHLILMMIPDRLDNPLGAQRKKR--- 297

Qy      282 TNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPY 341
      ||||| ||: |:|||| | ||||: ||||: ||||: |||| | ||||
Db      298 -----ALDTNYCFRNLEENCCVRPLYIDFRQDLGWKWVHEPKGYANFCSGPCPY 347

Qy      342 IWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCK 401
      : | || :| || || | || ||: |||| | || | :||| | ||||| :|:||
Db      348 LRSADTTHSSVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTAKVEQLSNMVVKSK 407

Qy      402 CS 403
      ||
Db      408 CS 409

```

RESULT 15

TGF3_MOUSE

```

ID      TGF3_MOUSE      STANDARD;      PRT;      410 AA.
AC      P17125;
DT      01-AUG-1990 (Rel. 15, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 3 precursor (TGF-beta 3).
GN      TGFβ3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90190650; PubMed=2628730;
RA      Miller D.A., Lee A., Matsui Y., Chen E.Y., Moses H.L., Derynck R.;
RT      "Complementary DNA cloning of the murine transforming growth
RT      factor-beta 3 (TGF beta 3) precursor and the comparative expression
RT      of TGF beta 3 and TGF beta 1 messenger RNA in murine embryos and
RT      adult tissues.";
RL      Mol. Endocrinol. 3:1926-1934(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91000714; PubMed=2206556;
RA      Denhez F., Lafyatis R., Kondaiah P., Roberts A.B., Sporn M.B.;
RT      "Cloning by polymerase chain reaction of a new mouse TGF-beta,
RT      mTGF-beta 3.";
RL      Growth Factors 3:139-146(1990).
CC      -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.
CC      -!- SUBUNIT: Homodimer; disulfide-linked.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

DR      EMBL; M32745; AAA40422.1; -.
DR      PIR; A41397; A41397.
DR      HSSP; P10600; 1TGJ.
DR      MGD; MGI:98727; Tgfb3.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFB.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFBETA.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal.
FT      SIGNAL          1          23          POTENTIAL.
FT      PROPEP          24         298          POTENTIAL.
FT      CHAIN           299        410          TRANSFORMING GROWTH FACTOR BETA 3.
FT      DISULFID        305        314          BY SIMILARITY.
FT      DISULFID        313        376          BY SIMILARITY.
FT      DISULFID        342        407          BY SIMILARITY.
FT      DISULFID        346        409          BY SIMILARITY.
FT      DISULFID        375        375          INTERCHAIN (BY SIMILARITY).
FT      CARBOHYD         72         72          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        133        133          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        140        140          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      SITE            259        261          CELL ATTACHMENT SITE (POTENTIAL).
SQ      SEQUENCE        410 AA;   46884 MW;   250F7048CA432BD6 CRC64;

```

QY	15	LLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RGQILSKRLASPPSQGDVP PGPLP	74
Db	7	LVLALLNLATISLSLSTCTTLDFGHIKKKRVEAIRGQILSKRLTSPPEPSVMT--HVP	64
QY	75	EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMVE---SGNQIYDKFKG	126
Db	65	YQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFDMIQGLAEHNELAVCPKG	124
QY	127	TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL----KLKVEQHVELYQKYSND----S	178
Db	125	ITSKVFR-FNVSSVEK--NGTNLFRAEFRVL RVPNPNSSKRTEQRIELFQILRPDEHIAK	180
QY	179	WRYL SNRLLAPSDSPEWL SFDVTGVVRQWL TRREAIEGFRLSAH-----SSSDSKDN-	230
Db	181	QRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRR ES NLGLEISIHCPCHTFQPNGDILENV	240
QY	231	--TLHVEINGFNS---GRRGDLATI HGM---NRPFL LLMATPLERAQHLHSSRHRRALDT	282

```

Db      241 HEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHHNPHLI LMMI PPHRLDSPGQGSQRK---- 296
Qy      283 NDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYI 342
          : ||||| : |:||| |||||:|||||:|||||:||| |||||:
Db      297 -----KRALDTNYCFRNLEENCCVRPLYIDFRQDLGWKWVHEPKGYANFCSGPCPYL 349
Qy      343 WSLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKC 402
          | || :| || ||| || |||:||||| ||||| :||| ||||| |||||:||| |||
Db      350 RSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCK 409
Qy      403 S 403
          |
Db      410 S 410

```

Search completed: October 28, 2003, 09:08:41
Job time : 9.07747 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 07:50:55 ; Search time 32.3099 Seconds
(without alignments)
3218.683 Million cell updates/sec

Title: US-10-017-372E-33
Perfect score: 2139
Sequence: 1 MAPSGRLRLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*

```

11: sp_rodent:*
12: sp_virus:*
13: sp_vertibrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	1817.5	85.0	390	6	Q9TUM8	Q9tum8	equus cabal
2	1738.5	81.3	368	11	Q8R4D9	Q8r4d9	sigmodon hi
3	854.5	39.9	379	13	Q8JHF5	Q8jhf5	sparus aura
4	850.5	39.8	379	13	Q8AXK8	Q8axk8	sparus aura
5	832.5	38.9	412	11	Q91YU7	Q91yu7	mus musculu
6	832	38.9	382	13	Q9PWA9	Q9pwa9	morone chry
7	826	38.6	414	11	Q91VP5	Q91vp5	mus musculu
8	810	37.9	382	13	O93449	O93449	oncorhynchu
9	780	36.5	376	13	Q9PTQ2	Q9ptq2	cyprinus ca
10	765	35.8	399	11	Q9ERB7	Q9erb7	mesocricetu
11	724.5	33.9	362	11	Q99K17	Q99k17	mus musculu
12	693.5	32.4	130	11	Q08714	Q08714	mesocricetu
13	678.5	31.7	361	13	Q98854	Q98854	cyprinus ca
14	676.5	31.6	124	6	Q95N80	Q95n80	canis famil
15	638	29.8	112	6	O02730	O02730	oryctolagus
16	605	28.3	255	11	Q921T1	Q921t1	mus musculu
17	586.5	27.4	127	6	Q9TV08	Q9tv08	canis famil
18	575	26.9	224	11	Q8CDZ9	Q8cdz9	mus musculu
19	563.5	26.3	200	13	Q90YF1	Q90yf1	pleuronecte
20	554	25.9	101	11	Q9R184	Q9r184	meriones un
21	461	21.6	179	13	Q90YF2	Q90yf2	pleuronecte
22	413	19.3	88	13	Q90YF5	Q90yf5	pleuronecte
23	402	18.8	88	13	Q90YF7	Q90yf7	oncorhynchu
24	397	18.6	88	13	Q90ZE7	Q90ze7	acipenser b
25	393	18.4	87	13	O42306	O42306	carassius a
26	383	17.9	91	6	Q9MYZ1	Q9myz1	capra hircu
27	376.5	17.6	309	4	Q8WV88	Q8wv88	homo sapien
28	373	17.4	86	6	Q28241	Q28241	cervus elap
29	358	16.7	81	6	Q9N1S3	Q9n1s3	capreolus c
30	340.5	15.9	375	13	Q8UWD8	Q8uwd8	columba liv
31	327.5	15.3	375	13	Q8AVB2	Q8avb2	coturnix co
32	325.5	15.2	375	13	Q8UWD7	Q8uwd7	coturnix ch
33	325.5	15.2	389	13	Q90YY0	Q90yy0	ictalurus p
34	320.5	15.0	375	13	Q98SP0	Q98sp0	gallus gall
35	320	15.0	87	13	Q8JHB6	Q8jhb6	scophthalmu
36	317.5	14.8	375	6	Q9GM97	Q9gm97	equus cabal
37	317.5	14.8	375	13	Q8UWE0	Q8uwe0	anas platyr
38	317	14.8	77	13	Q90YF8	Q90yf8	oncorhynchu
39	315.5	14.7	375	13	Q8UWD9	Q8uwd9	anser anser
40	313.5	14.7	375	6	Q8HY52	Q8hy52	lepus capen

41	310.5	14.5	375	6	Q95J86	Q95j86 macaca fasc
42	302	14.1	50	6	Q28240	Q28240 cervus elap
43	301	14.1	62	13	Q90ZJ7	Q90zj7 anguilla an
44	301	14.1	376	13	Q98TB4	Q98tb4 oreochromis
45	300.5	14.0	375	6	Q8WNS6	Q8wns6 bos taurus

ALIGNMENTS

RESULT 1

```

Q9TUM8
ID   Q9TUM8          PRELIMINARY;      PRT;   390 AA.
AC   Q9TUM8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transforming growth factor beta 1.
GN   TGFB1.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nixon A.J., Brower-Toland B.T., Sandell L.J.;
RT   "Molecular cloning of equine transforming growth factor beta 1 reveals
RT   a divergent nucleotide structure that encodes a novel bioactive
RT   peptide among mammalian species.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF175709; AAD49431.1; -.
DR   HSSP; P01137; 1KLA.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   InterPro; IPR003911; TGF_TGFb.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PRINTS; PR01423; TGFBETA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
SQ   SEQUENCE   390 AA;  43860 MW;  220FE40DFCCA6016 CRC64;

Query Match          85.0%;  Score 1817.5;  DB 6;  Length 390;
Best Local Similarity 86.4%;  Pred. No. 1e-152;
Matches 348;  Conservative 12;  Mismatches 30;  Indels 13;  Gaps 1;

```

Qy	1	MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMMVESGNQI	120
		: :	
Db	61	SPPSOGEVPPGPLPEAVLALYNSTRAOVAGESAETEPEPEADYYAKEVTRVLMEKENEI	120

Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		: : : : :	
Db	121	YKTVETGSHSIYMFNASELRAAVPDPMLLSRAELRLLRLKLSVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDTVGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGFN	240
		: : : :	
Db	181	YLSNRLLTPSDSPEWLSFDTVGVVRQWLSQGGAMEGLRLSAHCPCDSKDNTLRVGINGFS	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFS	300
Db	241	SSRRGDLATIDGMNRPFLLLMATPLERAQQLHSSRHR-----ALDTNYCSS	287
Qy	301	STEKNCCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	360
Db	288	STEKNCCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	347
Qy	361	PGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	403
Db	348	PGASAAPCCVPOVLEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 2

08R4D9

```

ID      Q8R4D9          PRELIMINARY;          PRT;          368 AA.
AC      Q8R4D9;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Transforming growth factor beta-1 protein (Fragment).
GN      TGFB1.
OS      Sigmodon hispidus (Hispid cotton rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC      Sigmodon.
OX      NCBI_TaxID=42415;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Blanco J.C., Pletneva L.M., Prince G.A.;
RT      "Cotton rat cytokines, chemokines, and interferons.";
RL      Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; AF480858; AAL87199.1; -.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
FT      NON_TER          1          1
SQ      SEQUENCE      368 AA;  41905 MW;  A5C91207B0468B4A CRC64;

```

Query Match 81.3%; Score 1738.5; DB 11; Length 368;
Best Local Similarity 85.3%; Pred. No. 9.8e-146;

Matches 325; Conservative 16; Mismatches 27; Indels 13; Gaps 1;

```

Qy      23 PGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGDVPPGGLPEAVLALYN 82
      |||
Db      1 PGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGEVPPGGLPEAVLALYN 60
      |||

Qy     83 STRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELRE 142
      |||
Db     61 STRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAIYDKTKDIPHSVYMFNTSDIRE 120
      |||

Qy    143 AVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTG 202
      |||
Db    121 AVPEPPLLSRAELRLQRFKSNVEQHVELYEKYSNNSWRYLGNRLLSPTDSPEWLSFDVTS 180
      |||

Qy    203 VVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFNSGRRGDLATIHGMNRPFLLLLMA 262
      |||
Db    181 VVRKWLNQGDGIQGFRLSAHCSCSKDNILHVEINGISPKRRGDLGTIHDNMRPFLLLLMA 240
      |||

Qy    263 TPLERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWK 322
      |||
Db    241 TPLERAQHLHSSRHRR-----ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWK 287
      |||

Qy    323 WIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYY 382
      |||
Db    288 WIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASASPCCVQALEPLPIVYY 347
      |||

Qy    383 VGRKPKVEQLSNMIVRSCKCS 403
      |||
Db    348 VGRKPKVEQLSNMIVRSCKCS 368
      |||

```

RESULT 3

Q8JHF5

```

ID   Q8JHF5          PRELIMINARY;          PRT;   379 AA.
AC   Q8JHF5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transforming growth factor beta 1.
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tafalla C., Aranguren R., Secombes C.J., Castrillo J.L., Novoa B.,
RA   Figueras A.;
RT   "Molecular characterization of sea bream (Sparus aurata) transforming
RT   growth factor beta1.";
RL   Fish and Shellfish Immunol. 0:0-0(2002).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF510084; AAN03842.1; -.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.

```


RN [1]
 RP SEQUENCE FROM N.A.
 RA Tafalla C., Novoa B., Aranguren R., Figueras A.;
 RT "Molecular cloning and characterization of sea bream (*Sparus aurata*)
 RT TGF beta 1.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF424703; AAN76665.1; -.
 SQ SEQUENCE 379 AA; 43485 MW; A306EC387F6DBA7C CRC64;

Query Match 39.8%; Score 850.5; DB 13; Length 379;
 Best Local Similarity 46.2%; Pred. No. 6e-67;
 Matches 187; Conservative 68; Mismatches 109; Indels 41; Gaps 15;

Qy	12	LLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI	RQILSKLRL-ASPPSQGDVPP	70
		: ::: : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : :	
Db	3	LVFLMFMVAYTVGK-VSGMSTCKTLDLEMVKKKRIEAI	RSQILSKLRLPTESPQAGD--E	59
Qy	71	GPLPEAVLALYNSTRDRVAGESVEPEPE-----	PEADYYAKEVTRVLMVESGNQIYDKFK	125
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :	
Db	60	EEIPSSLLSLYNSTKEMLKEQQTEVQTDIFTEXEE	EEYFA---TRVHKFNTTNPV-----	111
Qy	126	GTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKV---	EQHVELYQKYSNDSWRYL	182
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :	
Db	112	RTPQNMSMSFNISEIRRSIGDYRLLTAE LRML-IKAPTILDEQR	VELYQGLGT-SPRYL	169
Qy	183	SNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGF--	RLSAHSSSDSKDNTLHVEINGFN	240
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :	
Db	170	ASRFITNELRDKWLSFDVTETLQNWLGNDVQVFQLRLY	CDGRSSDVSTFSFGISGMT	229
Qy	241	SGRRGD LATIHGMNR--PFLLLMATPLERAQHLHSSR	HRRALDTNDYKDDDDKALDTNYC	298
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :	
Db	230	AG-RGDKAVLDDMTKQPPYILTMSIPKNVSSHL-TSR	KKRSTETKD-----TC	275
Qy	299	FSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANF	CLGPCPYIWSLDTQYSKVLALYNQ	358
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :	
Db	276	TAQTE-TCCVRSLYIDFRKDLGWKWIHKPTRYHANY	CMGSCTYIWNAENKYSQXLALYKH	334
Qy	359	HNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNM	IVRSCKCS	403
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :	
Db	335	HNPGASAPCCVPQALEPLPILYVGRQHKVEQLSNM	IVKSKCS	379

RESULT 5

Q91YU7

ID Q91YU7 PRELIMINARY; PRT; 412 AA.
 AC Q91YU7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transforming growth factor, beta 3.
 GN TGFB3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; BC014690; AAH14690.1; -.
 DR MGD; MGI:98727; Tgfb3.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR InterPro; IPR003911; TGF_TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbeta.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 SQ SEQUENCE 412 AA; 47144 MW; F3EB65D046DF32AD CRC64;

Query Match 38.9%; Score 832.5; DB 11; Length 412;
 Best Local Similarity 44.7%; Pred. No. 2.7e-65;
 Matches 188; Conservative 58; Mismatches 126; Indels 49; Gaps 13;

Qy	15	LLWLLVLT	TPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLASPPSQGDVPPG	PLP	74
		: :	:	: : :		:
Db	9	LVLALLNL	ATISLSLSTCTTLD	FGHIKKKRV	EAIRGQILSKLRLTSPPEPSVMT	--HVP 66
Qy	75	EAVLALYNSTR---	DRVAGESVE--PEPEPEADYYAKEVTRVLMVE---	SGNQIYDKFKG		126
			: :	: : : : :		
Db	67	YQVLALYNSTRELL	EEMHGEREEGCTQETSESEYYAKEIHKFDMIQGLAEHNELAVCPKG			126
Qy	127	TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---	KLKVEQHVELYQKYSND----	S		178
		: :	: :	: : :		
Db	127	ITSKVFR-FNVSSVEK---	NGTNLFRAEFRVLRVPNPSSKRTEQRIELFQILRPDEHIAK			182
Qy	179	WRYLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGFRLSAH-----	SSSDSKDN-			230
		: :	:	: : :	: :	
Db	183	QRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENV				242
Qy	231	--TLHVEINGFNS---	GRRGDLATI HGM--NRPFLLLMATPLERAQHLHSSRHRRALDT			282
		: : : : :		: :	:	
Db	243	HEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHHNPHLILMMIPPHRLDSPGQGSQRK----				298
Qy	283	NDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYI				342
		:	: :		:	
Db	299	-----KRALDTNYCFRNLEENCCVRPLYIDFRQDLGWKQVHEPKGYANFCGSPCPYL				351
Qy	343	WSLDTQYSKVLALYNQHNP	GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKC			402
		:	:	:		
Db	352	RSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCKC				411
Qy	403	S	403			
Db	412	S	412			

ID Q9PWA9 PRELIMINARY; PRT; 382 AA.
 AC Q9PWA9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transforming growth factor beta precursor.
 GN TGF-BETA.
 OS Morone chrysops x Morone saxatilis (white bass x striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OX NCBI_TaxID=45352;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=KIDNEY;
 RX MEDLINE=20394636; PubMed=10938723;
 RA Harms C.A., Kennedy-Stoskopf S., Horne W.A., Fuller F.J.,
 RA Tompkins W.A.F.;
 RT "Cloning and sequencing hybrid striped bass (Morone saxatilis x M.
 RT chrysops) transforming growth factor-beta (TGF-beta), and development
 RT of a reverse transcription quantitative competitive polymerase chain
 RT reaction (RT-qcPCR) assay to measure TGF-beta mRNA of teleost fish.";
 RL Fish Shellfish Immunol. 10:61-85(2000).
 CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE
 CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: HIGHER LEVELS FOUND IN MONONUCLEAR CELLS FROM
 CC PERIPHERAL BLOOD THAN IN SPLEEN OR ANTERIOR KIDNEY.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AF140363; AAD46997.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR InterPro; IPR003911; TGF_TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbeta.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP ? 270
 FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA.
 FT DISULFID 278 286 BY SIMILARITY.
 FT DISULFID 285 348 BY SIMILARITY.
 FT DISULFID 314 379 BY SIMILARITY.
 FT DISULFID 318 381 BY SIMILARITY.
 FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT SITE 235 237 CELL ATTACHMENT SITE (POTENTIAL) .
SQ SEQUENCE 382 AA; 43846 MW; 3124D8C34EA74D72 CRC64;

Query Match 38.9%; Score 832; DB 13; Length 382;
Best Local Similarity 45.9%; Pred. No. 2.7e-65;
Matches 187; Conservative 64; Mismatches 108; Indels 48; Gaps 15;

```
Qy      15 LLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRL--ASPPSQGDVPPG 71
      |: ::| | | :|:|||||:|:|:|:||||| | | | | | | | | | | | |
Db      6 LMLVVVYTVGN-VSGMSTCKTLDLEMVKKKRIEAIHQILSKLRLPKPEPDQAGDEEEI 64

Qy      72 PLPEAVLALYNSTRDRVAGESVEPE----PEPEADYYAKEVTRVLMVESGNQIYDKFKG 126
      | | :|:|||||: : : | : : : | :|:| | : : | | | |
Db      65 PTP--LLSLYNSTKEMLKEQQTEVQTDISTEQEEEEYFAKVLHKFNMTRKNN----- 114

Qy     127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKV--EQHVELYQKYSNDSWRYLSN 184
      | : | | | | :|:|:| : | | : | | | :| | : | | | | | | | | | |
Db     115 TDTTKKMFNISEIRESVGDYRLLTSAELRMLIKKTTIYDEQRVELYSGL-GDSPRYLAS 173

Qy     185 RLLAPSDSPEWLSFSDVTGVVRQWLTRREAIEGF--RLSAHSSSDSKDNTL-HVEINGFNS 241
      | : : : | | | | | : : | | : : | | : | | : | : | :
Db     174 RFITNKWKDKWLSFSDVTKTLDWLKGTDDQGGFQLRLFCENKVSAGETIFKFGISGIDP 233

Qy     242 GRRGDLATIHGMNR--PFLLLMATPLERAQHLHS--SRHRRALDTNDYKDDDDKALDTN 296
      | | | | : : : |:| | : | | : : | | :|:| :| |
Db     234 G-RGDTGPMQLLTQQPPYILTMSIP----QNISSPSTSRKKRSTETKD----- 276

Qy     297 YCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALY 356
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     277 VCTAQTE-TCCVRSLYIDFRKDLGWKWIHKPTGYNANYCMGSCTYIWNAENKYSQILALY 335

Qy     357 NQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 403
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     336 KHHNPASAAAPCCVPQALEPLPIVYVGRQHKVEQLSNMIVKSKCS 382
```

RESULT 7

Q91VP5

ID Q91VP5 PRELIMINARY; PRT; 414 AA.
AC Q91VP5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to transforming growth factor, beta 2.
GN TGFB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; BC011170; AAH11170.1; -.
DR MGD; MGI:98726; Tgfb2.

DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR InterPro; IPR003911; TGF_TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbBETA.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 SQ SEQUENCE 414 AA; 47588 MW; DB37A7C38881F286 CRC64;

Query Match 38.6%; Score 826; DB 11; Length 414;
 Best Local Similarity 43.7%; Pred. No. 1e-64;
 Matches 188; Conservative 63; Mismatches 121; Indels 58; Gaps 15;

Qy 12 LLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVP-P 70
 :| || | | | |||| |::| : ||||| ||||| ||||| :| ||| | | |
 Db 5 VLSTFLLHLVLP--VALSLSTCSTLDMDQFMRKRIEAIHQILSKLKLTSPP--DYPEP 60

Qy 71 GPLPEAVLALYNSTRD---RVAGESVEPEPE-PEADYYAKEVTRVLM---VESGNQIYD 122
 :| |::| ||||| : : : | | : : ||||| : : | : | |
 Db 61 DEVPPEVISIYNSTRDLLQEKASRRAAACERERSDEEYYAKEVYKIDMPSHLPSENAIPP 120

Qy 123 KFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKV-EQHVELYQ-----K 173
 | | : | : : : : : | : || | : | | : | |||||
 Db 121 TFY-RPYFRIVRFVDVSTMEKNASN---LVKAEFRVFRQLQNPKEARVAEQRIELYQILKSKD 176

Qy 174 YSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSS----- 225
 : : : || : : : ||||| || : || : : || : |
 Db 177 LTSPTQRYIDSKVVKTRAEGEWSFDVTDVQEWLHHKDRNLGFKISLHCPCTFVPSNN 236

Qy 226 ---DSKNTLHVE---INGFNSGRRGDLATIHGMNR-----PFLLLMATPLERAQHLHS 273
 :| | | :| : : || || : | |||| | | : |
 Db 237 YIIPNKSEELERFAGIDGTSTYASGDQKTIKSTRKKTSGKTPHLLMLLPSYRLESQQS 296

Qy 274 SRHRRALDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHAN 333
 || : : : ||| ||| : : ||| : | |||| : : ||||| ||||| : ||
 Db 297 SRRKK-----RALDAAAYCFRNVQDNCCLRPLYIDFKRDLGWKWIHEPKGYNAN 344

Qy 334 FCLGPCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLS 393
 || | ||| : || ||| : ||| : || || ||| : |||| | |||| | : ||| : || : |||||
 Db 345 FCAGACPYLWSSDTQHTKVLSTYNTINPEASASPCCVSQDLEPLTILYYIGNTPKIEQLS 404

Qy 394 NMIVRSCKCS 403
 |||| : |||||
 Db 405 NMIVKSKCS 414

RESULT 8

O93449

ID O93449 PRELIMINARY; PRT; 382 AA.

AC O93449; Q91217;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Transforming growth factor beta precursor.
 GN TGF-BETA OR TGF.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEUKOCYTE;
 RX MEDLINE=99242020; PubMed=10227481;
 RA Daniels G.D., Secombes C.J.;
 RT "Genomic organisation of rainbow trout, Oncorhynchus mykiss TGF-
 RT BETA.";
 RL Dev. Comp. Immunol. 23:139-147(1999).
 RN [2]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=LEUKOCYTE;
 RX MEDLINE=98390168; PubMed=9722928;
 RA Hardie L.J., Laing K.J., Daniels G.D., Grabowski P.S., Cunningham C.,
 RA Secombes C.J.;
 RT "Isolation of the first piscine transforming growth factor beta gene:
 RT analysis reveals tissue specific expression and a potential regulatory
 RT sequence in rainbow trout (Oncorhynchus mykiss).";
 RL Cytokine 10:555-563(1998).
 CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE
 CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BLOOD LEUKOCYTES, KIDNEY
 CC MACROPHAGES, BRAIN, GILL AND SPLEEN BUT NOT IN LIVER.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AJ007836; CAA07707.1; -.
 DR EMBL; X99303; CAA67685.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR InterPro; IPR003911; TGF_TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbeta.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 270
 FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA.
 FT DISULFID 278 286 BY SIMILARITY.
 FT DISULFID 285 348 BY SIMILARITY.
 FT DISULFID 314 379 BY SIMILARITY.
 FT DISULFID 318 381 BY SIMILARITY.
 FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 237 237 N -> D (IN REF. 2).
 FT CONFLICT 345 345 Q -> H (IN REF. 2).
 FT CONFLICT 371 372 LS -> VP (IN REF. 2).
 FT CONFLICT 377 377 K -> M (IN REF. 2).
 SQ SEQUENCE 382 AA; 44136 MW; 93BD4D3540084B92 CRC64;

Query Match 37.9%; Score 810; DB 13; Length 382;
 Best Local Similarity 46.8%; Pred. No. 2.4e-63;
 Matches 182; Conservative 58; Mismatches 105; Indels 44; Gaps 14;

Qy 30 LSTCKTIDMELVKRKRIEAIHQILSKLRLASPP--SQGDVPPGGLPEAVLALYNSTRD 86
 :|||::|:||||||||||||||||||||||| | :|| :| ::::| ||| :
 Db 23 MSTCKSLDLELVKRKRIEAIHQILSKLRLPKEPEIDQEGDTE--EVPASLMSIYNSTVE 80
 Qy 87 RVAGESVE-----PEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSEL 140
 :: | | : | || |:| ||| : | :| | : | : ||| ||| :
 Db 81 -LSEEQVHTYIPSTQDAEEEA-YFAKEVHKFNMQSENT-----SKHQI--LFNMSEM 129
 Qy 141 REAVPEPVLLSRAELRLL----RLKLVKEQHVELYQKYSNDSWRYLSNRL LAPSDSPEWL 196
 | : |||:|||||| | || :|||: | ||| : :: : | :
 Db 130 RSVLGTDRLLSQAELRLLIKNHGLLDDSEQRLELYRGV-GDKARYLKSHFVSKEWANRWV 188
 Qy 197 SFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFNSGRRGDLATI--HGMN 254
 ||||| : :|| :||:| :||:| | ||: | : :
 Db 189 SFDVTQTLNEWLQGAGEEQGFQLKLPCDCGKPMEEFRFKISGMNK-LRGNTETLAMKMPS 247
 Qy 255 RPFLLLMATPLERAQHLHSSRHRRALDNTDYKDDDDKALDTNYCFSSTEKNCCVRQLYID 314
 :| :|||: |:|| | ||| :| : | : |||:| ||| :
 Db 248 KPHILLMSLPVERHSQL-SSRKKRQTTTEE-----IC-SDKSESCCVRKLYID 293
 Qy 315 FRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAPCCVPQAL 374
 ||||| || |||:| ||| : : :||:| ||| ||||| ||||| |
 Db 294 FRKDLGWKWIHEPTGYFANYCIGPCTYIWNTENKYSQVLALYKHHNPGASAQPCVPQVL 353
 Qy 375 EPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
 |||||:|||||: ||||| |||||:||:| |
 Db 354 EPLPIIYYVGRQHKVEQLSNMIVKSCRCS 382

RESULT 9

Q9PTQ2

ID Q9PTQ2 PRELIMINARY; PRT; 376 AA.
 AC Q9PTQ2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transforming growth factor beta precursor.
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yin Z., Kuang J.;
 RT "Molecular cloning of carp transforming growth factor beta 1."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.


```

CC      -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE
CC      RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.
CC      -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; AF136947; AAF22573.1; -.
DR      HSSP; P01137; 1KLA.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      InterPro; IPR003911; TGF_TGFb.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR01423; TGFbeta.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFb; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal.
FT      SIGNAL          1          22          POTENTIAL.
FT      PROPEP          23          264
FT      CHAIN           265          376          TRANSFORMING GROWTH FACTOR BETA.
FT      DISULFID        272          280          BY SIMILARITY.
FT      DISULFID        308          373          BY SIMILARITY.
FT      DISULFID        312          375          BY SIMILARITY.
FT      DISULFID        341          341          INTERCHAIN (BY SIMILARITY).
FT      CARBOHYD         76           76          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        125          125          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        167          167          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      SITE            230          232          CELL ATTACHMENT SITE (POTENTIAL).
SQ      SEQUENCE        376 AA: 43329 MW: 7F7FC4DA58B69681 CRC64;

```

Query Match 36.5%; Score 780; DB 13; Length 376;
Best Local Similarity 44.0%; Pred. No. 1.1e-60;
Matches 179; Conservative 65; Mismatches 121; Indels 42; Gaps 15;

Qy	6	LRLLLPLLLPLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQ	65
		: : : : : : : : : : : : : : : : :	
Db	1	MRVESLLLALQCLLGFV--HYSGALSTCSPLDLELIKRKRIEAIRGQILSKLRLSKEPEV	58
Qy	66	GDVPPGP-LPEAVLALYNST---RDRVAGESVEPEPEP-EADYYAKEVTR--VLMVESGN	118
		: : ::: : : : : : : : : :	
Db	59	DEEKESQNI PAELISVYNSTVELNEEQAAPPEQPKEDPVEEEYYAKEVHKFTIKLMEKNP	118
Qy	119	QIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLK--VEQHVELYQKYSN	176
		: : : : : : : :	
Db	119	---DKF-----LWFNITDISQTLGLNRIISQVELRLLIITTFPDGSEQRLELYQVIGN	167
Qy	177	DSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEI	236
		: : : : : : : : : : :	
Db	168	KS-RYLESRFI--PNQRKWLSFDVTQTLKDWLQRSEAEQGQFLKMADNCDPQ-KTFQLKI	223
Qy	237	NGFNSGRRGDLATIH-GMNRPFLLMATPLERAQHLHSSRHRRALDTNDYKDDDDKALDT	295
		: : : : : : : :	
Db	224	PGLVL-VRGDTETLAVNMMPRPHILVMSLPLDGN---NSSKSRRKRQT-----ET	268
Qy	296	NYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLAL	355
		: : : : : : : : :	
Db	269	DOVCTDKSDGCCVRSLYIDFRKDLGWKWIHEPSGYANYCTGSCSEVWTSENKYSOVLAL	328

Qy 356 YNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKC 402
| | | | | | | | | | | | | | | | : | | | | | | | : | |
Db 329 YKHHNPGASAOPCRVPQVLNPLPIFYVVGROHKVEQLSNMIVKTCKC 375

RESULT 10

```

Q9ERB7
ID   Q9ERB7          PRELIMINARY;          PRT;   399 AA.
AC   Q9ERB7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Transforming growth factor-beta 2 (Fragment).
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ramesh G., Kondaiah P., Seshagiri P.B.;
RT   "Differential expression and selective localization of transforming
RT   growth factor-beta isoforms in the hamster uterus during estrous
RT   cycle.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AY007214; AAG02247.1; -.
DR   HSSP; P08112; 2TGI.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   InterPro; IPR003911; TGF_TGFb.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR01423; TGFbeta.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFb; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
FT   NON_TER      1          1
FT   NON_TER      399       399
SQ   SEQUENCE      399 AA;  46078 MW;  A6FF8E65EAFD5148 CRC64;

```

Query Match 35.8%; Score 765; DB 11; Length 399;
Best Local Similarity 42.5%; Pred. No. 2.5e-59;
Matches 177; Conservative 60; Mismatches 121; Indels 58; Gaps 15;

Qy	18	LLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLASPPSQGDVP-PGPLPEA	76
		: : : :	
Db	4	LLHLVP--VALSLSTCSTLDMDQFMRKRIEAIRGQILSKLKLTSPPE--DYEPDDEVPPE	59
Qy	77	VLALYNSTRD----RVAGESVEPEPE-PEADYYAKEVTRVLM---VESGNQIYDKFKGTP	128
		:: : : : : : : :	
Db	60	VISIYNSTRDLLQEASRRAAACERERSDEEYYAKEVYKIDMPSHFPSENAIPTTFY-RP	118
Qy	129	HSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKV-EQHVELYQ-----KYSNDSW	179
		: : : : : : : : : : : :	
Db	119	YFRIVREFDVSMMEKNASN---LVKAEFRVFRLLONPKARVAEORIELYOILKSKDLTSPTO	175

```

Qy      180 RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSSDSK 228
      ||: :::: ||||| || | :|| :: ||::| | : :|
Db      176 RYIDSKVVKTRAEGEWLSFDVTDVHVEWLHKKDRNLGFKISLHCPCTFVPFNNNIIPNK 235

Qy      229 DNTLHVE---INGFNSGRRGDLATIHGMNR-----PFLLLMATPLERAQHLHSSRHRRA 279
      | :| : | || : | ||| | | : | :| ::
Db      236 SEELEARFAGIDGTSQHSSGHQETIKSTRKKNSGKTPHLLMLLPSYRLESQQSNRRKK- 294

Qy      280 LDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPC 339
      :||| ||| : : |||: |||||::||| |||||: ||| | |
Db      295 -----RALDAAYCFRNVQDNCCLRPLYIDFKRDLGWKWIHEPKGYNANFCAGAC 343

Qy      340 PYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNM 395
      ||: || |||: |||: || | |||: ||| ||| | :||: | ||: |||||
Db      344 PYLWSSDTQHTKVLSTYNTINPEASAPCCVSHDLEPLTILYYIGNTPKIEQLSNM 399

```

RESULT 11

Q99K17

```

ID   Q99K17      PRELIMINARY;          PRT;   362 AA.
AC   Q99K17;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to transforming growth factor, beta 3 (Fragment).
GN   TGFB3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; BC005513; AAH05513.1; -.
DR   HSSP; P10600; 1TGJ.
DR   MGD; MGI:98727; Tgfb3.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   InterPro; IPR003911; TGF_TGFb.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PRINTS; PR01423; TGFbBETA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFb; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
FT   NON_TER      1      1
SQ   SEQUENCE      362 AA;  41486 MW;  0808E46180FDAE70 CRC64;

```

```

Query Match          33.9%;  Score 724.5;  DB 11;  Length 362;
Best Local Similarity 43.4%;  Pred. No. 8.5e-56;
Matches 164;  Conservative 52;  Mismatches 113;  Indels 49;  Gaps 13;

```

```

Qy      58 RLASPPSQGDVPPGGLPEAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVL 112

```

```

      |: ||| : :| ||||| : : || | : :||| : :
Db      2 RVGSPPEPSVMT--HVPYQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFD 59
QY      113 MVE---SGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL----KLKVE 165
      |: : | : : || : : | : : | ||| :| : : |
Db      60 MIQGLAEHNELAVCPKGITSKVFR-FNVSSVEK--NGTNLFRAEFRVLRVPNPSSKRTE 115
QY      166 QHVELYQKYSND----SWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSA 221
      | :||:| | :| : : | : : ||||| | :|| ||| : |
Db      116 QRIELFQILRPDEHIAKQRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISI 175
QY      222 H-----SSSDSKDN---TLHVEINGFNS---GRRGDLATIHGM---NRPFLLLMTATPL 265
      | : | : | : : | : : |||| : : | :|| |
Db      176 HCPCHTFQPNGDILENVHEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHHNPHLILMMIPP 235
QY      266 ERAQHLHSSRHRRALDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIH 325
      | : | : : ||||| : :||| | ||||| :||| :|
Db      236 HRLDSPGQGSQRK-----KRALDTNYCFRNLEENCCVRPLYIDFRQDLGWKQVH 284
QY      326 EPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGR 385
      ||||| :||| ||||| : ||| : || ||| || ||| :||| ||| :|||
Db      285 EPKGYANFCSGPCPYLRSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGR 344
QY      386 KPKVEQLSNMIVRSCKCS 403
      ||||| :|||
Db      345 TPKVEQLSNMVVKSCCKCS 362

```

RESULT 12

Q08714

```

ID      Q08714      PRELIMINARY;      PRT;      130 AA.
AC      Q08714; O70331;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Transforming growth factor beta 1 (TGF-beta 1) (Fragment).
GN      TGFB1.
OS      Mesocricetus auratus (Golden hamster).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC      Mesocricetus.
OX      NCBI_TaxID=10036;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LVG (SYR);
RX      MEDLINE=93304479; PubMed=8317544;
RA      Wong D.T., Donoff R.B., Yang J., Song B.Z., Matossian K., Nagura N.,
RA      Elovic A., McBride J., Gallagher G., Todd R.;
RT      "Sequential expression of transforming growth factors alpha and beta 1
RT      by eosinophils during cutaneous wound healing in the hamster.";
RL      Am. J. Pathol. 143:130-142(1993).
RN      [2]
RP      SEQUENCE OF 26-115 FROM N.A.
RC      STRAIN=SYRIAN; TISSUE=SPLEEN;
RX      MEDLINE=98234044; PubMed=9573100;
RA      Melby P.C., Tryon V.V., Chandrasekar B., Freeman G.L.;
RT      "Cloning of Syrian hamster (Mesocricetus auratus) cytokine cDNAs and

```

[illegible]

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEART;
 RX MEDLINE=97354301; PubMed=9210595;
 RA Sumathy K., Desai K.V., Kondaiah P.;
 RT "Isolation of transforming growth factor-beta2 cDNA from a fish,
 RT Cyprinus carpio by RT-PCR.";
 RL Gene 191:103-107(1997).
 CC -!- FUNCTION: TGF-BETA 2 HAS SUPPRESSIVE EFFECTS ON INTERLEUKIN-2
 CC DEPENDENT T-CELL GROWTH.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; U66874; AAB62983.1; -.
 DR HSSP; P08112; 2TGI.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR InterPro; IPR003911; TGF_TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR01423; TGFbeta.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein.
 FT NON_TER 1 1
 FT PROPEP <1 257
 FT CHAIN 258 361 TRANSFORMING GROWTH FACTOR BETA 2.
 FT DISULFID 264 273 BY SIMILARITY.
 FT DISULFID 272 335 BY SIMILARITY.
 FT DISULFID 334 334 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 361 361
 SQ SEQUENCE 361 AA; 41931 MW; 94D930FA970A3FD3 CRC64;

Query Match 31.7%; Score 678.5; DB 13; Length 361;
 Best Local Similarity 40.8%; Pred. No. 1e-51;
 Matches 155; Conservative 55; Mismatches 115; Indels 55; Gaps 14;

Qy 52 QILSKLRLASPPSQGDVPPGP--LPEAVLALYNSTRDRVAGESVEPEPEPE----ADYY 104
 ||| ||:|: || :: | | : ::|:||||| : :: | : :||
 Db 1 QILCKLKLSCPP---EIYPEPEEVSRIIAIYNSTRDLLQEKANERAATCERQRTGEEYY 57
 Qy 105 AKEVTRVLM---VESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL- 160
 |||| :: | : | | | : : | : : | : ||||: ||
 Db 58 AKEVHKIDMQPFYPAENVIPTKHY-NPYFRRLRFDVSSMEKNASN---LVKAELRIFRLQ 113
 Qy 161 --KLKV-EQHVELYQ-----KYSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRRE 212
 | :| || :||| : : : ||: :: : ||||| | :|| |:
 Db 114 NPKARVSEQRIELYQILGHKDLTSPTQRYIDSKVVRTRTEGEWLSFDVTEAVSEWLLHRD 173
 Qy 213 AIEGFRLSAHSSS-----DSKDNTLHVEINGFNSG--RRGDLATI---HGMNR 255

```

      ||::| |      :| | | : || |
Db      174 RNNGFKISLHCPCTFVPSNNYIIPNKSEELERFAGIDDSFVHGGDLKMFKKRRHSGQS 233

Qy      256 PFLLLMATPLERAQHLHSSRRRRALDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDF 315
      | |||| | | : | | ||:      :||| :|| : : ||::| |||||
Db      234 PHLLMLLPSYRLESQHKSHRQ-----KRALDAAFCFRNVQDNCCLRSLYIDF 281

Qy      316 RKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALE 375
      : ||||| ||||| : |||| | |||: || |||: | :| ||| | | |||: |||| | ||
Db      282 KKDLGWKWIHEPKGYNANFCAGACPYLWSADTQHSNILGLYNTINPEASAPCCVSDLE 341

Qy      376 PLPIVYYVGRKPKVEQLSNM 395
      || |::|: ||: |||||
Db      342 PLTILYYIGKTPKIEQLSNM 361

```

RESULT 14

Q95N80

```

ID      Q95N80      PRELIMINARY;      PRT;      124 AA.
AC      Q95N80;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Transforming growth factor beta 1 (Fragment).
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Fonfara S., Groene A., Baumgaertner W.;
RT      "Sequence of canine transforming growth factor beta 1 mRNA in DH82-
RT      cells.";
RL      Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; AF349538; AAK54072.1; -.
DR      InterPro; IPR001839; TGFb.
DR      Pfam; PF00019; TGF-beta; 1.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFb; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
FT      NON_TER      1      1
FT      NON_TER      124      124
SQ      SEQUENCE      124 AA; 14329 MW; 21D185218E5556DB CRC64;

```

```

Query Match      31.6%; Score 676.5; DB 6; Length 124;
Best Local Similarity 89.8%; Pred. No. 3.3e-52;
Matches 123; Conservative 0; Mismatches 1; Indels 13; Gaps 1;

```

```

Qy      264 PLERAQHLHSSRRRRALDTNDYKDDDDKALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKW 323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 PLERAQHLHSSRQRR-----ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKW 47

Qy      324 IHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYV 383
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      48 IHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYV 107

```

Qy 384 GRKPKVEQLSNMIVRSC 400
 |||||
 Db 108 GRKPKVEQLSNMIVRSC 124

RESULT 15

O02730

ID O02730 PRELIMINARY; PRT; 112 AA.
 AC O02730; O97501;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).
 GN TGFB1 OR TGF-BETA-1.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Taylor T.K., James E.R., McGonigle S., Yoho E.R.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE OF 2-99 FROM N.A.
 RA Inoue K., Kawabe Y., Kodama T.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AF000133; AAB53806.1; -.
 DR EMBL; AB020217; BAA36950.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein.
 FT NON_TER 1 1
 FT CHAIN 1 112 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 7 16 BY SIMILARITY.
 FT DISULFID 15 78 BY SIMILARITY.
 FT DISULFID 44 109 BY SIMILARITY.
 FT DISULFID 48 111 BY SIMILARITY.
 FT DISULFID 77 77 INTERCHAIN (BY SIMILARITY).
 FT CONFLICT 2 3 LD -> FS (IN REF. 2).
 FT CONFLICT 85 92 PLPIVYYV -> ATAHRVTTL (IN REF. 2).
 SQ SEQUENCE 112 AA; 12795 MW; 53C5B7D46355A6F3 CRC64;

Query Match 29.8%; Score 638; DB 6; Length 112;

Best Local Similarity 100.0%; Pred. No. 7.3e-49;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      292 ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 351
          |||
Db      1 ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 60

Qy      352 VLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 403
          |||
Db      61 VLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112
```

Search completed: October 28, 2003, 09:12:26
Job time : 33.3099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 10:47:27 ; Search time 5276.94 Seconds
(without alignments)
10489.161 Million cell updates/sec

Title: US-10-017-372E-34
Perfect score: 1353
Sequence: 1 gatctggtaccgagatggcg.....cgattaaagcggccgcgact 1353

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

```

28:  em_un:*
29:  em_vi:*
30:  em_htg_hum:*
31:  em_htg_inv:*
32:  em_htg_other:*
33:  em_htg_mus:*
34:  em_htg_pln:*
35:  em_htg_rod:*
36:  em_htg_mam:*
37:  em_htg_vrt:*
38:  em_sy:*
39:  em_htgo_hum:*
40:  em_htgo_mus:*
41:  em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	1236.4	91.4	3206	4	PIGTGFB1A	M23703 Sus scrofa
2	1224.6	90.5	1326	6	AX338213	AX338213 Sequence
3	1217	89.9	1750	4	GGTGFB1	X12373 Porcine mRN
4	1146.8	84.8	1605	4	SSTGFB1	Y00111 Porcine mRN
5	1125.4	83.2	2221	4	AF461808	AF461808 Sus scrof
6	1005.8	74.3	1369	4	DOGTGFB1A	L34956 Canine tran
7	994.8	73.5	2527	6	E00973	E00973 cDNA encodi
8	994	73.5	2537	6	A06669	A06669 Synthetic m
9	988	73.0	1173	4	OATGFB1	X76916 O.aries mRN
10	987.6	73.0	1780	9	BC000125	BC000125 Homo sapi
11	987.6	73.0	1780	9	BC001180	BC001180 Homo sapi
12	986.2	72.9	1561	9	AGMTGFB	M16658 Simian tran
13	983.4	72.7	1821	6	E03028	E03028 DNA encodin
14	983	72.7	1746	9	BC022242	BC022242 Homo sapi
15	981.8	72.6	1560	6	I06216	I06216 Sequence 2
16	981	72.5	2745	9	HSTGFB1	X02812 Human mRNA
17	979.8	72.4	1560	6	I08268	I08268 Sequence 2
18	962.4	71.1	1569	6	I06221	I06221 Sequence 3
19	957.4	70.8	1561	6	I08275	I08275 Sequence 3
20	942.8	69.7	1173	9	BT007245	BT007245 Homo sapi
21	942.8	69.7	1173	12	BT007866	BT007866 Synthetic
22	933.6	69.0	1173	4	ECRGFB1	X99438 E.caballus
23	933.4	69.0	1176	6	AX615127	AX615127 Sequence
24	931.8	68.9	1176	6	AX481432	AX481432 Sequence
25	931.8	68.9	1176	6	AX615128	AX615128 Sequence
26	927.2	68.5	1187	4	AF175709	AF175709 Equus cab
27	904.4	66.8	1597	10	AF191297	AF191297 Cavia por
28	876	64.7	1641	10	BC013738	BC013738 Mus muscu
29	862.2	63.7	1579	10	MUSTGFRNA	M13177 Mouse trans
30	862.2	63.7	2094	10	MMU009862	AJ009862 Mus muscu
31	842.6	62.3	1585	10	RNTGFB1	X52498 Rat mRNA fo
32	830.8	61.4	1117	4	BOVTGFB	M36271 Bovine tran
33	806.2	59.6	1125	10	AF480858	AF480858 Sigmodon

34	685.4	50.7	1675	6	I03310	I03310 Sequence 1
35	657.2	48.6	1376	6	AX528533	AX528533 Sequence
36	655.6	48.5	1389	6	AX528619	AX528619 Sequence
37	594.6	43.9	1352	6	AX528535	AX528535 Sequence
38	589.8	43.6	1350	6	AX528615	AX528615 Sequence
39	558.8	41.3	699	6	I05434	I05434 Sequence 4
40	366.4	27.1	489	6	AX455100	AX455100 Sequence
41	357.6	26.4	650	6	AX336646	AX336646 Sequence
42	357.6	26.4	650	9	HUMTGFB4	M38449 Homo sapien
43	357.6	26.4	862	6	I03312	I03312 Sequence 3
44	339.2	25.1	469	10	MATGFB1	X60296 M.auratus m
45	335	24.8	1256	5	CHKTGFB4	M31160 Gallus gall

ALIGNMENTS

RESULT 1

PIGTGFB1A

LOCUS PIGTGFB1A 3206 bp mRNA linear MAM 31-MAR-1995

DEFINITION Sus scrofa transforming growth factor beta-1 mRNA, complete cds.

ACCESSION M23703

VERSION M23703.1 GI:755044

KEYWORDS transforming growth factor-beta-1.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 3206)

AUTHORS Kondaiah,P., Van Obberghen-Schilling,E., Ludwig,R.L., Dhar,R., Sporn,M.B. and Roberts,A.B.

TITLE cDNA cloning of porcine transforming growth factor-beta 1 mRNAs. Evidence for alternate splicing and polyadenylation

JOURNAL J. Biol. Chem. 263 (34), 18313-18317 (1988)

MEDLINE 89054010

PUBMED 2461367

COMMENT On Apr 1, 1995 this sequence version replaced gi:341017.

Original source text: Sus scrofa (strain miniature swine) cDNA to mRNA.

FEATURES

source

Location/Qualifiers

1. .3206

/organism="Sus scrofa"

/mol_type="mRNA"

/strain="miniature swine"

/db_xref="taxon:9823"

/cell_type="peripheral blood lymphocyte"

gene

1. .3206

/gene="TGF-beta-1"

CDS

906. .2078

/gene="TGF-beta-1"

/codon_start=1

/product="transforming growth factor-beta-1"

/protein_id="AAA64616.1"

/db_xref="GI:755045"

/translation="MPPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK

RIEAIHQILSKLRLASPPSQGDVPPGPLPEAVIALYNSTRDRVAGESVEPEPEPEAD

YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRL

```

KLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
LSAHCSCDSKDNTHLVEINGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPASAAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS"
polyA_signal      3186.  .3191
                  /gene="TGF-beta-1"
polyA_site        3206
                  /gene="TGF-beta-1"
BASE COUNT       645 a   1041 c   924 g   596 t
ORIGIN

```

```

Query Match          91.4%;  Score 1236.4;  DB 4;  Length 3206;
Best Local Similarity 96.3%;  Pred. No. 1.5e-208;
Matches 1292;  Conservative 0;  Mismatches 11;  Indels 39;  Gaps 1;

```

```

Qy      10  CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      ||  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      901  CCCCCATGCGCGCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 960

Qy      70  TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 129
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      961  TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 1020

Qy     130  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1021  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 1080

Qy     190  TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 249
      |  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1081  TCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 1140

Qy     250  CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 309
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1141  CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 1200

Qy     310  CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 369
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1201  CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 1260

Qy     370  AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG 429
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1261  AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG 1320

Qy     430  AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1321  AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 1380

Qy     490  GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 549
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1381  GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 1440

Qy     550  GGCCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCCTTTG 609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1441  GGCCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCCTTTG 1500

Qy     610  ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC 669
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	1501	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	1560
Qy	670	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1561	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	1620
Qy	730	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	789
Db	1621	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	1680
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1681	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGA-	1739
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1740	-----GCCCTGGATACCAACTACTGCT	1761
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1762	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	1821
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1822	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1881
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1882	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1941
Qy	1090	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	1149
Db	1942	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	2001
Qy	1150	TCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	2002	TCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	2061
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC	1269
Db	2062	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC	2121
Qy	1270	CCCACCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCCACTTG	1329
Db	2122	CCCACCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCCACTTG	2181
Qy	1330	GGATCGATTAAAGCGGCCGCGA	1351
Db	2182	GGATCGATTAAAGGTGGAGAGA	2203

RESULT 2

AX338213

LOCUS	AX338213	1326 bp	DNA	linear	PAT 09-JAN-2002
DEFINITION	Sequence 1 from Patent WO0181404.				
ACCESSION	AX338213				
VERSION	AX338213.1 GI:18128750				

Qy	361	GCGGCAACCAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCA	420
Db	362	GCGGCAACCAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCA	421
Qy	421	ACACGTCGGAGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGC	480
Db	422	ACACGTCGGAGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGC	481
Qy	481	GCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCA	540
Db	482	GCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCA	541
Qy	541	ATGATTCTTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CT CACCGGAGTGGC	600
Db	542	ATGATTCTTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CT CACCGGAGTGGC	601
Qy	601	TGTCCTTTGATGTACACGGAGTTGTGCGGCAGTGGCTGACCCG CAGAGAGGCTATAGAGG	660
Db	602	TGTCCTTTGATGTACACGGAGTTGTGCGGCAGTGGCTGACCCG CAGAGAGGCTATAGAGG	661
Qy	661	GTTTTTCGCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAA	720
Db	662	GTTTTTCGCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAA	721
Qy	721	TTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATT CACGGCATGAACCGGC	780
Db	722	TTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATT CACGGCATGAACCGGC	781
Qy	781	CCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGC	840
Db	782	CCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGC	841
Qy	841	ACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCA	900
Db	842	ACCGCCGA-----GCCCTGGATACCA	862
Qy	901	ACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCC	960
Db	863	ACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCC	922
Qy	961	GGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCC	1020
Db	923	GGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCC	982
Qy	1021	TGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGT	1080
Db	983	TGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGT	1042
Qy	1081	ACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGC	1140
Db	1043	ACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGC	1102
Qy	1141	CACTGCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGA	1200
Db	1103	CACTGCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGA	1162


```

QY      1201 TCGTGC GTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGC 1260
          |||
Db      1163 TCGTGC GTTCCTGCAAGTGCAGCTGA-GCCCCGCCCCGCCACAGCCCCGCCACCCGGC 1221

QY      1261 AGGCCCCGCCCCACCCCGCCCGCCTACCGGGGCTGTATTTAAGGACATCGTGCCCCAA 1320
          |||
Db      1222 AGGCCCCGCCCCACCCCGCCCGCCTACCGGGGCTGTATTTAAGGACATCGTGCCCCAA 1281

QY      1321 GCCCACTTGGGATCGATTAAAGCGGCCCGCA 1351
          |||
Db      1282 GCCCAC-TGGGATCGATTAAAGGTGGAGAGA 1311

```

RESULT 3

GGTGFB1

LOCUS GGTGFB1 1750 bp mRNA linear MAM 27-MAR-1996

DEFINITION Porcine mRNA for transforming growth factor-beta 1.

ACCESSION X12373

VERSION X12373.1 GI:63808

KEYWORDS transforming growth factor-beta 1.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 1750)

AUTHORS Jakowlew,S.B., Dillard,P.J., Sporn,M.B. and Roberts,A.B.

TITLE Nucleotide sequence of chicken transforming growth factor-beta 1
(TGF-beta 1)

JOURNAL Nucleic Acids Res. 16 (17), 8730 (1988)

MEDLINE 88335639

PUBMED 3166520

REFERENCE 2 (bases 1 to 1750)

AUTHORS Jakowlew,S.B.

TITLE Direct Submission

JOURNAL Submitted (14-JUL-1988) Jakowlew S.B., National Institute of
health, National Cancer Institute, Laboratory of Chemoprevention,
Building 41, Room B902, Bethesda, Maryland 20892, USA

COMMENT The submitters believe that the chicken cDNA library was
contaminated with porcine cDNA, and that the sequence is infact
porcine TGF-beta-1. 27-MAR-1996.

FEATURES Location/Qualifiers

source 1..1750

/organism="Sus scrofa"

/mol_type="mRNA"

/strain="white leghorn"

/db_xref="taxon:9823"

/clone="pTGFB-ChX119"

/cell_type="chondrocyte"

gene 1..1750

/gene="TGF-beta 1"

5'UTR 1..446

/gene="TGF-beta 1"

CDS 447..1622

/gene="TGF-beta 1"

/codon_start=1

/product="transforming growth factor"

/protein_id="CAA30933.1"

```

/db_xref="GI:63809"
/db_xref="SWISS-PROT:P07200"
/translation="MPPSGPGLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
RIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNDSWGYLSNRLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFR
LSAHCSCDSKDNTLHVEINAGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSR
HRRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD
TQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
mat_peptide 1285. .1619
3'UTR /gene="TGF-beta 1"
1620. .1750
/gene="TGF-beta 1"
BASE COUNT 325 a 627 c 479 g 319 t
ORIGIN

Query Match 89.9%; Score 1217; DB 4; Length 1750;
Best Local Similarity 95.8%; Pred. No. 4.2e-205;
Matches 1288; Conservative 0; Mismatches 15; Indels 42; Gaps 2;

QY 10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 CCCCCATGCCGCCTTCGGGGCCTGGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 501

QY 70 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG 129
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG 561

QY 130 AGCTGGTGAAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 AGCTGGTGAAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 621

QY 190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 249
| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 TCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 681

QY 250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC 309
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC 741

QY 310 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC 369
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 742 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC 801

QY 370 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 802 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 861

QY 430 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 862 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 921

QY 490 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 549
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 922 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 981

QY 550 GCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGA CTACCGGAGTGGCTGTCCTTTG 609

```

Db	982	 GGGGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	1041
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	669
Db	1042	 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	1101
Qy	670	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAAC--G	726
Db	1102	 TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGCAG	1161
Qy	727	GGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC	786
Db	1162	 GGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC	1221
Qy	787	TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCC	846
Db	1222	 TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCC	1281
Qy	847	GAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACT	906
Db	1282	 GA-----GCCCTGGATACCAACTACT	1302
Qy	907	GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGG	966
Db	1303	 GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGG	1362
Qy	967	ACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGC	1026
Db	1363	 ACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGC	1422
Qy	1027	CCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACC	1086
Db	1423	 CCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACC	1482
Qy	1087	AGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGC	1146
Db	1483	 AGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGC	1542
Qy	1147	CCATCGTGTAATACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1206
Db	1543	 CCATCGTGTAATACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1602
Qy	1207	GTTCTGTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCC	1266
Db	1603	 GTTCTGTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCC	1662
Qy	1267	GGCCCCACCCCGCCCGCCTACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCAC	1326
Db	1663	 GGCCCCACCCCGCCCGCCTACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCAC	1722
Qy	1327	TTGGGATCGATTAAAGCGGCCGCGA	1351
Db	1723	 TTGGGATCGATTAAAGGTGGAGAGA	1747

Qy	130	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	189
Db	519	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	578
Qy	190	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	249
Db	579	TCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	638
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	309
Db	639	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	698
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC	369
Db	699	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC	758
Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	759	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	818
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	819	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	878
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	549
Db	879	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	938
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	609
Db	939	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	998
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGCTATAGAGGGTTTTCGCC	669
Db	999	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGCTATAGAGGGTTTTCGCC	1058
Qy	670	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1059	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	1118
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGC	789
Db	1119	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGC	1178
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1179	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	1237
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1238	-----GCCCTGGATACCAACTACTGCT	1259
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1260	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	1319
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029

```

Db      1320  |||||TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1379
Qy      1030  GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1089
Db      1380  |||||GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1439
Qy      1090  ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCCA 1149
Db      1440  ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCCA 1499
Qy      1150  TCGTGTA|CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1209
Db      1500  TCGTGTA|CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1559
Qy      1210  CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCAC 1255
Db      1560  CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCAC 1605

```

RESULT 5

AF461808

LOCUS AF461808 2221 bp mRNA linear MAM 03-JAN-2002

DEFINITION Sus scrofa transforming growth factor beta 1 (TGFB1) mRNA, complete cds.

ACCESSION AF461808

VERSION AF461808.1 GI:18042250

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Polymorphism in the porcine transforming growth factor beta 1 gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Direct Submission

JOURNAL Submitted (20-DEC-2001) Institute of Animal Breeding Science, University of Bonn, Endenicher Allee 15, Bonn 53115, Germany

FEATURES Location/Qualifiers

source 1. .2221
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/chromosome="6"

gene 1. .2221
/gene="TGFB1"

CDS 1. .1173
/gene="TGFB1"
/note="cytokine"
/codon_start=1
/product="transforming growth factor beta 1"
/protein_id="AAL57902.1"
/db_xref="GI:18042251"
/translation="MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK"


```

Db      601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCTCAGT 660
Qy      675 GCCCACTGTTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 734
      |||
Db      661 GCCCACTGTTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
Qy      735 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 794
      |||
Db      721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
Qy      795 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 854
      |||
Db      781 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGA----- 834
Qy      855 GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC 914
      |||
Db      835 -----GCCCTGGATACCAACTACTGCTTCAGC 861
Qy      915 TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC 974
      |||
Db      862 TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC 921
Qy      975 TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC 1034
      |||
Db      922 TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC 981
Qy      1035 TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAAC 1094
      |||
Db      982 TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAAC 1041
Qy      1095 CCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTG 1154
      |||
Db      1042 CCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTG 1101
Qy      1155 TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGC 1214
      |||
Db      1102 TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGC 1161
Qy      1215 AAGTGCAGCTGAGGCCCCGCCCC 1237
      |||
Db      1162 AAGTGCAGCTGAGGCCCCAAGCC 1184

```

RESULT 6

DOGTGFB1A

LOCUS DOGTGFB1A 1369 bp mRNA linear MAM 30-OCT-1994

DEFINITION Canine transforming growth factor-beta 1 (TGFB1) mRNA, complete cds.

ACCESSION L34956

VERSION L34956.1 GI:516071

KEYWORDS homologue; transforming growth factor-beta 1.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 1369)

AUTHORS Manning,A.M., Auchampach,J.A., Drong,R.F. and Slightom,J.L.

Db	233	TCTCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGTGCCCGTGCCCCGAGGCCGTGCTGG	292
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	309
Db	293	CCCTCTACAACAGCACCCGCGACCGGGTGGCGGGGAGAGCGCCGAGCCGGAGCCCGAGC	352
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACC	369
Db	353	CCGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAACACCAACA	412
Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	413	AAATCTATGAGAAAAGTCAAGAAAAGTCCGCACAGCATATATATGCTCTTCAACACATCAG	472
Qy	430	AGCTCCGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	473	AGCTCCGAGAAGCAGTGCTGAGCCCGTCTTGCTCTCCCGGGCAGAGTTGCGCCTGCTGA	532
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT	549
Db	533	GGCTCAAGTTAAAAGCGGAGCAGCATGTGGAGCTGTACCAGAAATATAGCAATGATTCTCT	592
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTCAACCGAGTGGCTGTCCTTTG	609
Db	593	GGCGCTACCTCAGCAACCGGCTGCTGGCGCCAGCGACACGCCAGAATGGCTGTCCTTTG	652
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGCTATAGAGGGTTTTTCGCC	669
Db	653	ATGTCACTGGAGTCGTGAGGCAGTGGCTGAGCCATGGAGGGGAAGTCGAGGGCTTTTCGCC	712
Qy	670	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	713	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTGCAAGTAGACATTAACGGGT	772
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	789
Db	773	TCAGTTCAGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGACCTTCCTGC	832
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCGAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	833	TCCTCATGGCCACCCCACTGGAGAGGGCCCGAGCACCTGCACAGCTCCCGGCAGCGCCG--	890
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	891	-----GGCCCTGGACACCAACTACTGCT	913
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	914	TCAGCTCCACGGAGAAGAACTGCTGCGTCCGGCAGCTCTACATTGACTTCCGCAAGGATC	973
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	974	TGGGCTGGAAGTGGATCCATGAGCCCAAGGGTTACCACGCTAACTTCTGCCTGGGGCCCT	1033
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACACAGC	1089

```

Db      1034  GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1093
Qy      1090  ACAACCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA 1149
        |||||
Db      1094  ACAACCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA 1153
Qy      1150  TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1209
        |||||
Db      1154  TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTGGAACATGATCGTGCGCT 1213
Qy      1210  CCTGCAAGTGCAGCTGAGGCCCGCCCG-CCCACAGCCCCGCCACCCGGCAGGCCCGG 1268
        |||||
Db      1214  CCTGCAAGTGCAGCTGAGGCCCGCCCGTCCGGCAGGCCCGCCACCCGGCAGNCCGG 1273
Qy      1269  CCCACCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGC 1322
        |||||
Db      1274  CCCCGCCCCCGCCCGCTGCGCC-GGGCTGTATTTAAGGACACCGCGCCCCAGGC 1326

```

RESULT 7

E00973

LOCUS E00973 2527 bp RNA linear PAT 29-SEP-1997

DEFINITION cDNA encoding human TGF-beta.

ACCESSION E00973

VERSION E00973.1 GI:2169234

KEYWORDS JP 1986219395-A/1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2527)

AUTHORS Riku,M.A.D.D. and Debitsudo,B.G.

TITLE NUCLEIC ACID CODED WITH TGF-BETA AND ITS USE

JOURNAL Patent: JP 1986219395-A 1 29-SEP-1986;

GENENTECH INC

COMMENT

OS human

PN JP 1986219395-A/1

PD 29-SEP-1986

PF 20-MAR-1986 JP 1986064661

PR 22-MAR-1985 US 85 715142

PI RIKU MAIKERU ANDORE DERINKU, DEBITSUDO BANNOOMAN GETSUDERU PC

C12P21/00,C12N1/00,C12N5/00,C12N15/00//C12Q1/68,(C12P21/00, PC
C12R1:91),

PC (C12N1/00,C12R1:19),(C12N5/00,C12R1:91),(C12N15/00,C12R1:91);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC *source: tissue_type=placenta and glyoblastoma; FH Key
Location/Qualifiers

FH

FT CDS 842. .2014

FT /product='pre TGF-beta'

FT mat_peptide 1676. .2011

FT /product='TGF-beta'

FT 5'UTR 1. .841

FT 3'UTR 2015. .2537

```

FT      stem_loop      37. .113
FT      stem_loop      2015. .2100
FT      polyA_site      2514. .2519.
FEATURES                      Location/Qualifiers
    source                1. .2527
                           /organism="Homo sapiens"
                           /mol_type="genomic RNA"
                           /db_xref="taxon:9606"
BASE COUNT      472 a      888 c      735 g      432 t
ORIGIN

```

```

Query Match      73.5%; Score 994.8; DB 6; Length 2527;
Best Local Similarity 85.9%; Pred. No. 6.9e-166;
Matches 1144; Conservative 0; Mismatches 147; Indels 41; Gaps 2;

```

```

Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      || ||| |||| || ||||||||||| ||||||||||| ||||||||||| ||
Db      837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||||||||||||||| ||||||||||| ||||| ||||||||||| |||||||
Db      897 TGGTGCTGACGCCTGGCCCGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy      130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||||||||||||||| ||||||||||| ||||||||||| ||||||||||| ||||
Db      957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy      190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 249
      | ||||||||||||||| ||||||||||| ||||||||||| ||||||| || ||
Db      1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCAGGCCGTGCTCG 1076

Qy      250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 309
      | || ||||||||| ||||||||||| ||||||||| ||||| |||||||||||
Db      1077 CCCTCTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy      310 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 369
      | ||||| ||||||||||| ||||||||||| ||||||||||| ||||||| ||||
Db      1137 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG 1196

Qy      370 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
      ||||||||| || ||||| || ||||||| ||||||| ||||||| || ||
Db      1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATGTTCTTCAACACATCAG 1256

Qy      430 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
      ||||| ||||||||| || ||||| || ||||||| ||||||| |||||||
Db      1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA 1316

Qy      490 GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 549
      ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||||
Db      1317 GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT 1376

Qy      550 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGCGTGTCTTTT 609
      ||||| ||||||||||| ||||||||||| ||||| ||||| || ||||| ||
Db      1377 GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTT 1436

Qy      610 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGGCTATAGAGGGTTTTCGCC 669
      ||||||||||| ||||||||||| || ||| ||| || ||||| |||||||

```

Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	670	TCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	789
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1616
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCT	1817
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1818	GCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877
Qy	1090	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149
Db	1878	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1937
Qy	1150	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1938	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGGAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCCGCCCGCCCCACAGCCCCGCCACCCGGCAGGCCCCGGC	1269
Db	1998	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCC--CCGCCCCGCCCGGCCCCACCCCGCCCC	2055
Qy	1270	CCCACCCCGCCCGCCTCACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCCACTTG	1329
Db	2056	GCCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAGCCCACTTG	2115
Qy	1330	GGATCGATTAAA	1341
Db	2116	GGGCCCCATTAA	2127

Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	549
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	1376
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	609
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGCGTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	669
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGCTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	670	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	789
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGC	1616
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCGAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCCGAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	910	TCAGTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1698	TCAGTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	970	TGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCT	1817
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1818	GCCCCATACATTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877
Qy	1090	ACAACCCGGGCGCGTCCGGCGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	1149
Db	1878	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCA	1937
Qy	1150	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1938	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC	1269

Query Match 73.0%; Score 988; DB 4; Length 1173;
Best Local Similarity 89.8%; Pred. No. 1.2e-164;
Matches 1088; Conservative 0; Mismatches 85; Indels 39; Gaps 1;

[illegible]

```

Db      721 TCCGGCCGCCGGGGTGACCTCGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
Qy      795 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 854
      |||||
Db      781 ATGGCCACCCCTCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGA----- 834
Qy      855 GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC 914
      |||||
Db      835 -----GCCCTGGACACCAACTACTGCTTCAGC 861
Qy      915 TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC 974
      |||||
Db      862 TCCACAGAAAAGAACTGCTGTGTTTCGTGAGCTCTACATTGACTTCCGGAAGGACCTGGGC 921
Qy      975 TGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC 1034
      |||||
Db      922 TGAAGTGGATTTCACGAACCCAAGGGCTACACGCCAATTTCTGCCTGGGGCCCTGTCCC 981
Qy      1035 TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAAC 1094
      |||||
Db      982 TACATCTGGAGCCTGGACACACAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCACAAAC 1041
Qy      1095 CCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCGGCGCTGGAGCCACTGCCCATCGTG 1154
      |||||
Db      1042 CCGGGCGCATCGGCGGCGCCGTGCTGCGTGCCCTCAGGCGCTGGAACCCCTGCCCATCGTG 1101
Qy      1155 TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC 1214
      |||||
Db      1102 TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGTTGTCCAACATGATCGTGCGCTCCTGC 1161
Qy      1215 AAGTGCAGCTGA 1226
      |||||
Db      1162 AAGTGCAGCTGA 1173

```

RESULT 10

BC000125

LOCUS BC000125 1780 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, Similar to transforming growth factor, beta 1, clone MGC:3119 IMAGE:3351664, mRNA, complete cds.

ACCESSION BC000125

VERSION BC000125.1 GI:12652748

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1780)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Qy	130	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATTCTGTCCAAGCTTCGGC	189
Db	562	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCTGTCCAAGCTTCGGC	621
Qy	190	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG	249
Db	622	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCAGGCCGTGCTCG	681
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTGAAACCGGAGCCCGAGC	309
Db	682	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAAACCGGAGCCCGAGC	741
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACC	369
Db	742	CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACG	801
Qy	370	AAATCTATGATAAAATTC AAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	802	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	861
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	862	AGCTCCGAGAAGCGGTACCTGAACCCGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	921
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	549
Db	922	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	981
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	609
Db	982	GGCGATACTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1041
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGGCTATAGAGGGTTTTCGCC	669
Db	1042	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1101
Qy	670	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1102	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1161
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	789
Db	1162	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1221
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1222	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1280
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1281	-----GCCCTGGACACCAACTATTGCT	1302
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1303	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGGAAGGACC	1362

Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1363	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT	1422
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1423	GCCCCACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1482
Qy	1090	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149
Db	1483	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1542
Qy	1150	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1543	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1602
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC	1269
Db	1603	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGGCAGGCCCCGGCCCCAC	1662
Qy	1270	CCCACCCCCGCCCCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1321
Db	1663	CCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG	1722
Qy	1322	CCCACTTGGGATCGATTAAA	1341
Db	1723	CCCACCTGGGGCCCCATTAA	1742

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: e Column: 10

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 37097.

```
FEATURES             Location/Qualifiers
    source             1..1780
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="MGC:2323 IMAGE:3356605"
                        /tissue_type="Eye, retinoblastoma"
                        /clone_lib="NIH_MGC_16"
                        /lab_host="DH10B-R"
                        /note="Vector: pOTB7"
    CDS                447..1619
                        /codon_start=1
                        /product="Similar to transforming growth factor, beta 1"
                        /protein_id="AAH01180.1"
                        /db_xref="GI:12654683"
                        /translation="MPPSGRLRLLLLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK
RIEAIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNNSWRYLSNRLAPSDSPEWLSFDVTGVVRQWLRSRGGEIEGFR
LSAHCSCDSRDNTLQVDINGFTTGRRGDLATIHGMRPFLLLMATPLERAQHLQSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"
BASE COUNT          341 a    648 c    493 g    298 t
ORIGIN
```

Query Match 73.0%; Score 987.6; DB 9; Length 1780;
Best Local Similarity 85.4%; Pred. No. 1.4e-164;
Matches 1144; Conservative 0; Mismatches 149; Indels 47; Gaps 2;

```
Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      || ||| ||| || ||||| ||| ||| ||||| ||||| ||||| |||
Db      442 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTACCGCTGCTGTGGCTAC 501

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      502 TGGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTATCCACCTGCAAGACTATCGACATGG 561

Qy     130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      562 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 621

Qy     190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCGCGCCGCTGCCTGAGGCAGTACTGG 249
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db	622	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCC GCCCGCCCGCTGCCCGAGGCCGTGCTCGC	681
Qy	250	CTCTTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCCGAGC	309
Db	682	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCCGAGC	741
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCA CCCCGCTGCTAATGGTGGAAAGCGGCAACC	369
Db	742	CTGAGGCCGACTACTACGCCAAGGAGGTCA CCCCGCTGCTAATGGTGGAAACCCACAACG	801
Qy	370	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG	429
Db	802	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	861
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	862	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	921
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	549
Db	922	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	981
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGA CTACCGGAGTGGCTGTCCTTTG	609
Db	982	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1041
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	669
Db	1042	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1101
Qy	670	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAAACGGGT	729
Db	1102	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1161
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCCTGC	789
Db	1162	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATT CATGGCATGAACCGGCCCTTCCTGC	1221
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCC CAGCACCTGCACAGCTCCCGCACCGCCGAG	849
Db	1222	TTCTCATGGCCACCCCGCTGGAGAGGGCC CAGCATCTGCAAAGCTCCCGGCACCGCCA-	1280
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1281	-----GCCCTGGACACCAACTATTGCT	1302
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1303	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1362
Qy	970	TGGGCTGGAAGTGGATT CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1363	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCA ACTTCTGCCTCGGGCCCT	1422
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1423	GCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1482

[illegible]

RESULT 12

AGMTGFB

LOCUS	AGMTGFB	1561 bp	mRNA	linear	PRI	27-APR-1993
-------	---------	---------	------	--------	-----	-------------

DEFINITION Simian transforming growth factor-beta (TGF) mRNA, complete cds.

ACCESSION M16658

VERSION M16658.1 GI:176552

KEYWORDS growth factor; transforming growth factor-beta.

SOURCE Cercopithecus aethiops (African green monkey)

ORGANISM *Cercopithecus aethiops*

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Cercopithecus.

REFERENCE 1 (bases 1 to 1561)

AUTHORS Sharples, K., Plowman, G.D., Rose, T.M., Twardzik, D.R. and
 Purchio, A.F.

TITLE	Cloning and sequence analysis of simian transforming growth factor-beta cDNA
-------	--

JOURNAL DNA 6 (3), 239-244 (1987)

MEDLINE 87246074

PUBMED 3474130

COMMENT Original source text: African green monkey cells (cell line BSC-40), cDNA to mRNA, clone pTGF-beta-2.

FEATURES	Location/Qualifiers
----------	---------------------

```
source      1.  .1561
```

/organism="Cercopithecus aethiops"

```
/mol type="mRNA"
```

```
/db xref="taxon:9534"
```

CDS 262.1434

```
/note="transforming growth factor-beta precursor"
```

/codon start=1

```
/protein id="AAA35369.1"
```

```
/db xref="GI:176553"
```

/translation="MPPSGLRILPLLLPLLWLLVITPSRPAAGLSTCKTIDMELVKRK

RIETIRGOILLSKIRLASPPSOGEVPPGPIPEAVIALYNSTRDBRVAGESAEPEPEPEAD

RYETTRGQQLSKRLKASFPFSQGEVFFGFLPEAVLALINSTRDRVAGESAEFEFEFEAD
YYAKEVTRVLMVETHNEIYDKFKOSTHSIYMEFNTSELBEAVPEPVLISRAELRLRL


```

KLKVEQHVELYQKYSNNSWRYLSNRL LAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFR
LSAHCSCDSKDNLTQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
sig_peptide      262. .1095
                  /note="transforming growth factor-beta signal peptide"
mat_peptide      1096. .1431
                  /product="transforming growth factor-beta"
BASE COUNT      301 a      547 c      446 g      267 t
ORIGIN

```

```

Query Match          72.9%;  Score 986.2;  DB 9;  Length 1561;
Best Local Similarity 85.8%;  Pred. No. 2.4e-164;
Matches 1150;  Conservative 0;  Mismatches 143;  Indels 48;  Gaps 3;

```

```

Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGGCTGCTGTGGCTGC 69
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      257 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCGGCTGCTGCTACCGCTGCTGTGGCTAC 316

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      317 TGGTGCTGACGCCTAGCCGGCCGGCCGGCCGAGACTATCCACCTGCAAGACTATCGACATGG 376

Qy      130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      377 AGCTGGTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 436

Qy      190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 249
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      437 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCCAGGCCGTGCTCG 496

Qy      250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTGCAACCGGAGCCCGAGC 309
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      497 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCCGAGCCCGAAC 556

Qy      310 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC 369
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      557 CGGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACG 616

Qy      370 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      617 AAATCTATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAG 676

Qy      430 AGCTCCGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      677 AGCTCCGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 736

Qy      490 GGCTCAAGTTAAAAGTGGAGCAGCAGTGGAGCTATACCAGAAATACAGCAATGATTCT 549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      737 GGCTCAAGTTAAAAGTGGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCT 796

Qy      550 GCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCCTTTG 609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      797 GCGGATACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGAGTGGTTGTCTTTG 856

Qy      610 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC 669
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	857	ATGTCACCGGAGTTGTGTCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCC	916
Qy	670	TCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	917	TTAGCGCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGT	976
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	789
Db	977	TCACTACCGGCCGCCGAGGTGACCTGGCCACAATTTCATGGCATGAACCGGCCCTTCTCTGC	1036
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1037	TTCTCATGGCCACCCCACTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-	1095
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1096	-----GCCCTGGACACCAACTACTGCT	1117
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1118	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACC	1177
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTCTGCCTGGGGCCCT	1029
Db	1178	TGGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTGGGGCCCT	1237
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1238	GTCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1297
Qy	1090	ACAACCCGGGCGCGTCTGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149
Db	1298	ATAACCCGGGCGCCTCTGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1357
Qy	1150	TCGTGTACTACGTGGGCCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1358	TCGTGTACTACGTGGGCCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1417
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC	1269
Db	1418	CCTGCAAAATGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGCCCGC	1477
Qy	1270	CCCACCCCGCCCGCCT-----CACCAGGGCTGTATTTAAGGACA-TCGTGCCCCAA	1320
Db	1478	CCCACCCCAACCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAA	1537
Qy	1321	GCCCACTTGGGATCGATTAAA	1341
Db	1538	GCCCACTTGGGGGCCCAATTAA	1558

Qy	130	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATTCTGTCCAAGCTTCGGC	189
Db	627	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATCCTGTCCAAGCTGCGGC	686
Qy	190	TTGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG	249
Db	687	TCGCCAGCCCCCGAGCCAGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG	746
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGC	309
Db	747	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	806
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC	369
Db	807	CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG	866
Qy	370	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	867	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	926
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	549
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	1046
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTTG	609
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1106
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGGCTATAGAGGGTTTTCGCC	669
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTCGCC	1166
Qy	670	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1167	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1226
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	789
Db	1227	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGC	1286
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCACTACTGCT	909
Db	1346	-----GCCCTGGACACCAACTATTGCT	1367
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1368	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGGAAGGACC	1427

[illegible]

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 27 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

```
FEATURES             Location/Qualifiers
     source            1. .1746
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="MGC:22008 IMAGE:4399762"
                        /tissue_type="Duodenum, adenocarcinoma"
                        /clone_lib="NIH_MGC_88"
                        /lab_host="DH10B"
                        /note="Vector: pCMV-SPORT6"
     CDS               370. .1542
                        /codon_start=1
                        /product="Unknown (protein for MGC:22008)"
                        /protein_id="AAH22242.1"
                        /db_xref="GI:18490116"
                        /translation="MPPSGRLRLLLLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
RIEAI RGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLRL
KLKVEQHVELYQKYSNNSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFR
LSAHCSCDSRDNTLQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"
BASE COUNT          376 a    612 c    472 g    286 t
ORIGIN
```

Query Match 72.7%; Score 983; DB 9; Length 1746;
Best Local Similarity 85.6%; Pred. No. 8.8e-164;
Matches 1148; Conservative 0; Mismatches 145; Indels 48; Gaps 3;

```
Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      365 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTAC 424

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      425 TGGTGCTGACGCCTGGCCGGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG 484

Qy     130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     485 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 544

Qy     190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 249
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     545 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 604

Qy     250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 309
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     605 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 664
```

QY	310	CAGAGGCGGACTACTACGCCAAGGAGGTCA	369
Db	665	CTGAGGCCGACTACTACGCCAAGGAGGTCA	724
QY	370	AAATCTATGATAAAATTC AAGGGCACCC	429
Db	725	AAATCTATGACAAGTTC AAGCAGAGTACAC	784
QY	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTAT	489
Db	785	AGCTCCGAGAAGCGGTACCTGAACCCGTGT	844
QY	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGG	549
Db	845	GGCTCAAGTTAAAAGTGGAGCAGCACGTGG	904
QY	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCC	609
Db	905	GGCGGATACCTCAGCAACCGGCTGCTGGCACC	964
QY	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGAC	669
Db	965	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAG	1024
QY	670	TCAGTGCCCACTGTTCTGTGACAGCAAAGATA	729
Db	1025	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATA	1084
QY	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACC	789
Db	1085	TCACTACCGGCCGCCGAGGTGACCTGGCCACC	1144
QY	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGC	849
Db	1145	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGC	1203
QY	850	CCCTGGATACCAACGACTACAAGGATGACGAC	909
Db	1204	-----GCCCTGGACACCAACTATTGCT	1225
QY	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	969
Db	1226	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	1285
QY	970	TGGGCTGGAAGTGGATTGATGAACCCAAGGG	1029
Db	1286	TCGGCTGGAAGTGGATCCACGAGCCCAAGGG	1345
QY	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTAC	1089
Db	1346	GCCCCCTACATTTGGAGCCTGGACACGCAGT	1405
QY	1090	ACAACCCGGGCGCGTGGCGGCGCCGTGCTGC	1149
Db	1406	ATAACCCGGGCGCCTGGCGGCGCCGTGCTGC	1465

```

Qy      1150 TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1209
          |||
Db      1466 TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT 1525

Qy      1210 CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC 1269
          |||
Db      1526 CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCCGCCCGCCCCGGCAGGCCCCGGCCCCAC 1585

Qy      1270 CCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAA 1320
          |||
Db      1586 CCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCCGTGCCCCAA 1645

Qy      1321 GCCCACTTGGGATCGATTAAA 1341
          |||
Db      1646 GCCCACCTGGGGCCCCATTAA 1666

```

RESULT 15

I06216

LOCUS I06216 1560 bp DNA linear PAT 02-DEC-1994

DEFINITION Sequence 2 from Patent EP 0293785.

ACCESSION I06216

VERSION I06216.1 GI:590649

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1560)

AUTHORS Purchio,A.F., Gentry,L. and Twardzik,D.

TITLE Cloning and expression of simian transforming growth factor-SS1

JOURNAL Patent: EP 0293785-A2 2 07-DEC-1988;

FEATURES Location/Qualifiers

source 1. .1560

/organism="unknown"

BASE COUNT 301 a 547 c 442 g 267 t 3 others

ORIGIN

Query Match 72.6%; Score 981.8; DB 6; Length 1560;
 Best Local Similarity 85.7%; Pred. No. 1.5e-163;
 Matches 1145; Conservative 0; Mismatches 143; Indels 48; Gaps 3;

```

Qy      15 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 74
          |||
Db      261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320

Qy      75 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 134
          |||
Db      321 CTGACGCCTAGCCGGCCGGCCCGGAGACTATCCACCTGCAAGACTATCGACATCGAGCTG 380

Qy      135 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 194
          |||
Db      381 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 440

Qy      195 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 254
          |||
Db      441 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 500

```


Qy	255	TACAAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG	314
Db	501	TACAAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCGGAGCCCGAACCGGAG	560
Qy	315	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	374
Db	561	GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	375	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC	434
Db	621	TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	435	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	494
Db	681	CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	495	AAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	554
Db	741	AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	800
Qy	555	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGATGTC	614
Db	801	TACCTCAGCAACCGGCTGCTGGCGCCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	615	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTCAGT	674
Db	861	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCTTAGC	920
Qy	675	GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	734
Db	921	GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	735	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	794
Db	981	ACCGGCCGCCGAGGTGACCTGGCCACAATTATGGCATGAACCGGCCCTTCTGCTTCTC	1040
Qy	795	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	854
Db	1041	ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-----	1094
Qy	855	GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC	914
Db	1095	-----GCCCTGGACACCAACTACTGCTTCAGC	1121
Qy	915	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC	974
Db	1122	TCCACGGAGAAGAACTGCTNCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGC	1181
Qy	975	TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1034
Db	1182	TGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1241
Qy	1035	TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAAC	1094
Db	1242	TACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAAC	1301
Qy	1095	CCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1154

```
Database :      N_Geneseq_19Jun03:*
1:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
```

```

6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	1224.6	90.5	1326	24	AAD22696	Porcine transformi
2	994.8	73.5	2527	25	ABQ76674	Androgen receptor
3	994	73.5	2537	7	AAN60972	Sequence encoding
4	994	73.5	2537	11	AAQ03301	cDNA encoding huma
5	994	73.5	2537	11	AAQ02814	Sequence of pre-TG
6	994	73.5	2537	17	AAT15720	Pre-transforming g
7	992.4	73.3	2537	15	AAQ56923	Human pre-TGF-beta
8	989.2	73.1	2537	19	AAV52933	Human pre-transfor
9	986.2	72.9	1561	11	AAQ03268	Simian transformin
10	984.6	72.8	2742	22	AAI58342	Human polynucleoti
11	983.8	72.7	1559	13	AAQ20289	Sequence encoding
12	983.4	72.7	1821	12	AAQ13392	Human pro-TGF-beta
13	982.8	72.6	1560	9	AAN81084	Coding sequence of
14	982.8	72.6	1560	11	AAQ03508	Simian Transformin
15	981	72.5	2745	16	AAT05876	cDNA encoding tran
16	981	72.5	2745	22	AAH28216	Nucleotide sequenc
17	973	71.9	1303	11	AAQ09317	Monkey transformin
18	972	71.8	4105	15	AAQ55624	TGFbeta1 5'-UTR-CD
19	970	71.7	1571	11	AAQ03269	Human transforming
20	965	71.3	1569	9	AAN81085	Coding sequence of
21	961.8	71.1	1569	11	AAQ03509	Human Transforming
22	955.8	70.6	1561	11	AAQ04908	Sequence encoding
23	952.6	70.4	1561	13	AAQ29177	TGF-beta 1/beta 2
24	933.4	69.0	1176	25	ABV75391	TGFB1 Arg25Pro pol
25	931.8	68.9	1176	24	ABZ35738	Human TGF beta 1 p
26	931.8	68.9	1176	24	ABX09981	Human TGFbeta1 DNA
27	931.8	68.9	1176	24	ABV78162	Human TGF beta 1 D

	28	931.8	68.9	1176	24	ABL91703	Human polynucleoti
	29	931.8	68.9	1176	25	ABV75392	TGFB1 Arg25Pro pol
	30	911.2	67.3	1565	13	AAQ29178	TGF-beta 1. Homo
c	31	808.2	59.7	2765	22	AAI60128	Human polynucleoti
	32	791.4	58.5	2208	13	AAQ20291	Sequence encoding
	33	789.8	58.4	2206	11	AAQ03510	Human Transforming
	34	789.8	58.4	2207	11	AAQ03511	Hybrid transformin
	35	788.2	58.3	2207	11	AAQ05127	Human TGF-Betal/TG
	36	776.8	57.4	2217	10	AAN90768	Sequence of human
	37	770.6	57.0	2200	16	AAT04115	Simian-human hybri
c	38	700	51.7	2773	23	AAS84421	DNA encoding novel
	39	682	50.4	834	12	AAQ12192	Sequence encoding
	40	657.2	48.6	1376	24	ABK90341	DNA encoding LAP-m
	41	655.6	48.5	1389	24	ABK90344	DNA encoding LAP-h
	42	591.4	43.7	1352	24	ABK90342	DNA encoding mIFNB
	43	589.8	43.6	1350	24	ABK90343	DNA encoding huIFN
	44	538	39.8	875	23	AAS70979	DNA encoding novel
	45	366.4	27.1	489	24	ABL99528	Target canine gene

ALIGNMENTS

RESULT 1

AAD22696

ID AAD22696 standard; cDNA; 1326 BP.

XX

AC AAD22696;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-betal) cDNA.

XX

KW Porcine; transforming growth factor beta 1; TGF-betal; gene therapy;
 KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;
 KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; ss.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT CDS 16..1188

FT /*tag= a

FT /product= "Porcine TGF-betal mutant protein"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12980.

XX

PR 20-APR-2000; 2000US-199014P.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Strober W, Nakamura K, Kitani A, Fuss IJ;

XX

DR WPI; 2002-026155/03.
DR P-PSDB; AAE13596.

XX

PT Composition for treating autoimmune diseases e.g. inflammatory bowel
PT disease in humans, comprises vector containing transforming growth
PT factor-beta under the control of inducible promoter -

XX

PS Claim 1; Fig 1; 78pp; English.

XX

CC The invention relates to a composition containing a vector comprising a
CC gene encoding a regulatory transcription factor under the control of a
CC promoter encoding a transforming growth factor-beta (TGF-beta). The
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2
CC or TGF-beta3, its variants or homologues, by transfecting a cell which
CC is part of a host suspected of having an autoimmune disease, especially
CC inflammatory bowel disease (IBD), under conditions such that the
CC polypeptide encoded by the nucleic acid sequence in the vector is
CC expressed. The vector is delivered using a delivery system. The delivery
CC of the vector results in substantial elimination of symptoms of the
CC autoimmune disease and increased production of IL-10 by the host. The
CC composition is useful for treating various diseases with an autoimmune
CC component such as multiple sclerosis, rheumatoid arthritis, systemic
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis
CC and psoriasis, and also for assaying the expression of a gene in a cell.
CC The vector is further useful for screening of the effect of test
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.
CC The present sequence is a cDNA encoding porcine TGF-beta1 mutant.

XX

SQ Sequence 1326 BP; 263 A; 438 C; 392 G; 233 T; 0 other;

Query Match 90.5%; Score 1224.6; DB 24; Length 1326;
Best Local Similarity 96.3%; Pred. No. 3.7e-241;
Matches 1301; Conservative 0; Mismatches 9; Indels 41; Gaps 3;

Qy	1	GATCTGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGC	60
Db	2	GATCTGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGC	61
Qy	61	TGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCA	120
Db	62	TGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCA	121
Qy	121	TCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCA	180
Db	122	TCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCA	181
Qy	181	AGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGG	240
Db	182	AGCTTCGGCTCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGG	241
Qy	241	CAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGG	300
Db	242	CCGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGG	301
Qy	301	AGCCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGA	360
Db	302	AGCCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGA	361

Qy	361	GCGGCAACCAAATCTATGATAAAATTC AAGGGCACCCCCACAGCTTATATATGCTGTTCA	420
Db	362	GCGGCAACCAAATCTATGATAAAATTC AAGGGCACCCCCACAGCTTATATATGCTGTTCA	421
Qy	421	ACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGC	480
Db	422	ACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGC	481
Qy	481	GCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCA	540
Db	482	GCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCA	541
Qy	541	ATGATTCC TGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCCGGAGTGGC	600
Db	542	ATGATTCC TGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCCGGAGTGGC	601
Qy	601	TGTCCTTTGATGTG CACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGG	660
Db	602	TGTCCTTTGATGTG CACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGG	661
Qy	661	GTTTTTCGCCTCAGTGGCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAA	720
Db	662	GTTTTTCGCCTCAGTGGCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAA	721
Qy	721	TTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGC	780
Db	722	TTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGC	781
Qy	781	CCTTCCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGC	840
Db	782	CCTTCCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGC	841
Qy	841	ACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCA	900
Db	842	ACCGCCGA-----GCCCTGGATACCA	862
Qy	901	ACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCC	960
Db	863	ACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCC	922
Qy	961	GGAAGGACCTGGGCTGGAAGTGGATT CATGAACCCAAGGGCTACCATGCCAATTTCTGCC	1020
Db	923	GGAAGGACCTGGGCTGGAAGTGGATT CATGAACCCAAGGGCTACCATGCCAATTTCTGCC	982
Qy	1021	TGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGT	1080
Db	983	TGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGT	1042
Qy	1081	ACAACCAGCACAACCCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGC	1140
Db	1043	ACAACCAGCACAACCCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGC	1102
Qy	1141	CACTGCCCATCGTGTA CTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGA	1200
Db	1103	CACTGCCCATCGTGTA CTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGA	1162

```

Qy      1201 TCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCACAGCCCCGCCCACCCGGC 1260
          |||
Db      1163 TCGTGCGTTCCTGCAAGTGCAGCTGA-GCCCCGCCCCGCCCACAGCCCCGCCCACCCGGC 1221

Qy      1261 AGGCCCCGCCCCACCCCCGCCCCGCTCACCAGGGGCTGTATTTAAGGACATCGTGCCCCAA 1320
          |||
Db      1222 AGGCCCCGCCCCACCCCCGCCCCGCTCACCAGGGGCTGTATTTAAGGACATCGTGCCCCAA 1281

Qy      1321 GCCCACTTGGGATCGATTAAAGCGGCCGCGA 1351
          |||
Db      1282 GCCCAC-TGGGATCGATTAAAGGTGGAGAGA 1311

```

RESULT 2

ABQ76674

ID ABQ76674 standard; DNA; 2527 BP.

XX

AC ABQ76674;

XX

DT 26-MAR-2003 (first entry)

XX

DE Androgen receptor signalling pathway-associated DNA E00973.

XX

KW Androgen receptor; transactivation; modulator; Smad3; Smad4; Akt; TGF-B;
 KW signal transduction pathway; transforming growth factor-B; phosphatase;
 KW tensin; cytostatic; antiproliferative; cellular proliferation; cancer;
 KW E00973; ds.

XX

OS Synthetic.

XX

PN WO200282081-A2.

XX

PD 17-OCT-2002.

XX

PF 05-APR-2002; 2002WO-US11086.

XX

PR 06-APR-2001; 2001US-282266P.

PR 13-MAR-2002; 2002US-365060P.

XX

PA (UYRP) UNIV ROCHESTER.

XX

PI Chang C;

XX

DR WPI; 2003-046871/04.

XX

PT Modulating androgen receptor activity, by administering a compound that
 PT modulates receptor activity, inhibits receptor-signal transduction
 PT pathway/receptor-coactivator interaction or changes amount or receptor
 PT -

XX

PS Disclosure; Page 225-226; 302pp; English.

XX

CC This invention describes a novel method for modulating androgen receptor
 CC activity or androgen receptor-mediated transactivation activity in a
 CC cell. The method involves administering a compound which causes
 CC modulation of the androgen receptors activity and the inhibition of
 CC interaction between the receptor and a protein involved in a signal

CC transduction pathway. The compound also inhibits the interaction between
 CC the androgen receptor and a protein selected from Smad3, Smad4, Akt,
 CC transforming growth factor (TGF)-B and phosphatase and tensin homologues
 CC deleted on chromosome 10 (PTEN) or their fragments. The compounds of the
 CC invention have cytostatic and antiproliferative activity. The obtained
 CC composition is useful for treating any disease, where uncontrolled
 CC proliferation or cellular proliferation occurs such as cancer, e.g.
 CC prostate cancer. This sequence represents the androgen receptor
 CC transactivation signalling pathway modulator E00973 described in
 CC the method of the invention.

XX

SQ Sequence 2527 BP; 472 A; 888 C; 735 G; 432 T; 0 other;

Query Match 73.5%; Score 994.8; DB 25; Length 2527;
 Best Local Similarity 85.9%; Pred. No. 4e-194;
 Matches 1144; Conservative 0; Mismatches 147; Indels 41; Gaps 2;

Qy	10	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGCTGCTGTGGCTGC	69
Db	837	CCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCGCTGCTGCTACCGCTGCTGTGGCTAC	896
Qy	70	TAGTGCTGACGCCTGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGG	129
Db	897	TGGTGCTGACGCCTGGCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG	956
Qy	130	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC	189
Db	957	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
Qy	190	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG	249
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	309
Db	1077	CCCTCTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC	369
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	490	GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	549
Db	1317	GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	1376
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCCTTTG	609
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436


```

AC      AAN60972;
XX
DT      31-OCT-2002   (updated)
DT      28-OCT-1991   (first entry)
XX
DE      Sequence encoding preTGF-beta.
XX
KW      Transforming growth factor beta; cancer; wound healing.
XX
OS      Unidentified.
XX
FH      Key           Location/Qualifiers
FT      misc_structure 37..113
FT                               /*tag= a
FT                               /note= "Sequence can form stable hairpin loops"
FT      CDS           842..2014
FT                               /*tag= b
FT      mat_peptide   1676..2011
FT                               /*tag= c
XX
PN      EP200341-A.
XX
PD      10-DEC-1986.
XX
PF      21-MAR-1986;   86EP-0302112.
XX
PR      22-MAR-1985;   85US-0715142.
PR      13-MAR-1987;   87US-0025423.
XX
PA      (GETH ) GENENTECH INC.
XX
PI      Derynck RMA;
XX
DR      WPI; 1986-326875/50.
DR      P-PSDB; AAP61468.
XX
PT      TGF-beta prodn. from transformed hosts - useful esp. for treating
PT      wounds (J6 2/9/86).
XX
PS      Disclosure; Fig 1b; 26pp; English.
XX
CC      The gene product is known to stimulate cell proliferation and
CC      inhibit anchorage-dependent growth of a variety of human cancer cell
CC      lines, it is esp. useful in treatment of burns and the promotion of
CC      surface and internal wound healing. TGF-beta may be expressed from a
CC      transformed CHO cell line.
CC      (Updated on 31-OCT-2002 to add missing OS field.)
XX
SQ      Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match          73.5%;   Score 994;   DB 7;   Length 2537;
Best Local Similarity 85.7%;   Pred. No. 5.8e-194;
Matches 1148;   Conservative    0;   Mismatches 145;   Indels    47;   Gaps      2;

Qy      10  CCGAGATGGCGCCTTCGGGGCTGCGGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      837  CCCCCATGCCGCCCTCCGGGGCTGCGGGCTGCTGCCGCTGCTGTACCGCTGCTGTGGCTAC 896

```

Qy	70	TAGTGTGACGCCTGGCCGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGG	129
Db	897	TGGTGCTGACGCCTGGCCCGCCGGCCGGCGGGACTATCCACCTGCAAGACTATCGACATGG	956
Qy	130	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	189
Db	957	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
Qy	190	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	249
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCCAGGCCGTGCTCG	1076
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	309
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGCCCGGGAGAGTGACAGAACCGGAGCCCGAGC	1136
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	369
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	370	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGTGTTC AACACGTCGG	429
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	1316
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	549
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1376
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGA CTACCGGAGTG GCTGTCCTTTG	609
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTG GTTATCTTTTG	1436
Qy	610	ATGTCAACGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	669
Db	1437	ATGTCAACGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	670	TCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGGT	1556
Qy	730	TCAATTCTGGCCGCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCCTGC	789
Db	1557	TCACTACCGGCCCGGAGGTGACCTGGCCACCATTTCATGGCATGAACCGGCCCTTCCTGC	1616
Qy	790	TCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1617	TTCTCATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1676	-----GCCCTGGACACCAACTATTGCT	1699

Qy	910 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCCGGAAGGACC	969
Db	1698 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCCGCAAGGACC	1757
Qy	970 TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTCTGCCTGGGGCCCT	1029
Db	1758 TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT	1817
Qy	1030 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGCTGTGTACAACCAGC	1089
Db	1818 GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGCCCTGTACAACCAGC	1877
Qy	1090 ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149
Db	1878 ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1937
Qy	1150 TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1938 TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1210 CCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC	1269
Db	1998 CCTGCAAGTGCAGCTGAGGTCCCGCCCGCCCCGCCCGCCCCGGCAGGCCCCGGCCCCAC	2057
Qy	1270 CCCACCCCCGCCCGCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1321
Db	2058 CCCGCCCCGCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG	2117
Qy	1322 CCCACTTGGGATCGATTAAA	1341
Db	2118 CCCACCTGGGGCCCCATTAA	2137

```

FT          /*tag= c
FT          /note="stable hairpin loops"
FT  misc_feature  2015..2100
FT          /*tag= d
FT          /note="G-C rich sequence
FT          and a downstream TATA-like sequence"
XX
PN  US4886747-A.
XX
PD  12-DEC-1989.
XX
PF  13-MAR-1987; 87US-0025423.
XX
PR  13-MAR-1987; 87US-0025423.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Derynck RMA, Goeddel DV;
XX
DR  WPI; 1990-051338/07.
DR  P-PSDB; AAR05258.
XX
PT  Nucleic acid encoding transforming growth factor-beta -
PT  cloned into expression vectors for expression in eukaryotic host
PT  cells for therapeutic use
XX
PS  Disclosure; Fig 1b; 28pp; English.
XX
CC  It was obtained by an analysis of several overlapping cDNAs and gene
CC  fragments, leading to the detn. of a continuous sequence corresp. to the
CC  TGF-beta-1 precursor mRNA. It is useful in constructing vectors that
CC  encode biologically active transforming growth factor (TGF-beta),
CC  operably linked to DNA that encodes a secretory leader (SL). It, or a
CC  nucleic acid capable of hybridising with it, can also be labelled and
CC  used in diagnostic assays for DNA or mRNA encoding TGF-beta or related
CC  proteins.
CC  (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ  Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match          73.5%; Score 994; DB 11; Length 2537;
Best Local Similarity 85.7%; Pred. No. 5.8e-194;
Matches 1148; Conservative 0; Mismatches 145; Indels 47; Gaps 2;

Qy          10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
           || ||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy          70 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 129
           | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          897 TGGTGCTGACGCCTGGCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy          130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGC 189
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy          190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 249

```

Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	309
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC	369
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG	1196
Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	549
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT	1376
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCCTTTG	609
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	669
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	670	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	789
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTCTGC	1616
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1817
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089

```

Db      1818  GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1877
Qy      1090  ACAACCCGGGCGCGTTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA 1149
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1878  ATAACCCGGGCGCCTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCA 1937
Qy      1150  TCGTGTA TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1209
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1938  TCGTGTA TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT 1997
Qy      1210  CCTGCAAGTG CAGCTGAGGCCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC 1269
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1998  CCTGCAAGTG CAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGGCAGGCCCCGGCCCCAC 2057
Qy      1270  CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG 1321
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2058  CCCGCCCCGCCCGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG 2117
Qy      1322  CCCACTTGGGATCGATTAAA 1341
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2118  CCCACCTGGGGCCCCATTAA 2137

```

RESULT 5

AAQ02814

ID AAQ02814 standard; cDNA; 2537 BP.

XX

AC AAQ02814;

XX

DT 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)

DT 31-MAY-1989 (first entry)

XX

DE Sequence of pre-TGF-beta1 cDNA.

XX

KW Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth inhibition.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	842..2011
FT		/*tag= a
FT		/label=pre-TGF beta 1
FT	CDS	1677..2011
FT		/*tag= b
FT		/label=mature TGF-beta 1
FT	GC_signal	2015..2092
FT		/*tag= c
FT	misc_feature	2093..2099
FT		/*tag= d
FT		/label=TATA-like sequence
FT	stem_loop	37..113
FT		/*tag= e
FT	misc_feature	863..911
FT		/*tag= f
FT		/label=hydrophobic domain

XX
 PN WO8912101-A.
 XX
 PD 14-DEC-1989.
 XX
 PF 08-JUN-1988; 88WO-US01945.
 XX
 PR 08-JUN-1988; 88WO-US01945.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Dernyck RMA, Goeddel DV;
 XX
 DR WPI; 1990-007474/01.
 DR P-PSDB; AAR04034.
 XX
 PT Nucleotide sequence encoding transforming growth factor beta-3 -used as a
 PT probe, or to produce TGF beta 3, for growth inhibition of certain normal
 PT and neoplastic cells, eg A549.
 XX
 PS Disclosure; Fig. 1b; 61pp; English.
 XX
 CC Sequence encodes the 390 amino acid (AA) precursor transforming growth
 CC factor-beta 1 (pre-TGF-beta 1) polypeptide. The 5' untranslated region of
 CC the TGF-beta 1 mRNA is 841 bases long, is purine rich and has a region of
 CC potential secondary structure. The TATA-like sequence in the 3' untrans-
 CC lated region of the gene is presumably a polyadenylation signal. Mature
 CC TGF-beta 1 comprises the C-terminal 112 AA's of pre-TGF-beta 1 and is
 CC cleaved at the Arg-Arg dipeptide preceding its NH2 terminus. The nucleic
 CC acid encoding the second subtype of TGF-beta (TGF-beta 3) is useful as a
 CC probe or to produce TGF-beta 3 for inhibition of growth of normal and
 CC neoplastic cells.
 CC (Updated on 31-OCT-2002 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 73.5%; Score 994; DB 11; Length 2537;
 Best Local Similarity 85.7%; Pred. No. 5.8e-194;
 Matches 1148; Conservative 0; Mismatches 145; Indels 47; Gaps 2;

Qy 10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
 || ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896
 Qy 70 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 129
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 897 TGGTGCTGACGCCTGGCCCGCCGGCCGGCCGACTATCCACCTGCAAGACTATCGACATGG 956
 Qy 130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 189
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016
 Qy 190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 249
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG 1076

Qy	250	CTCTTTTACAACAGTACCCGCGACCCGGGTAGCCCGGGGAAAGTGTGCGAACCCGGAGCCCGAGC	309
Db	1077	CCCTGTACAACAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGCAGAACCCGGAGCCCGAGC	1136
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	369
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	370	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	549
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1376
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	609
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTATCTTTTG	1436
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGCTATAGAGGGTTTTCGCC	669
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	670	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGAATCAACGGGT	1556
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCT	789
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTGCT	1616
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	970	TGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1758	TGGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1817
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1818	GCCCCATACATTTGGAGCCTGGACACGCGAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877


```

FT                                     may be important for transcription efficiency"
FT repeat_unit      2019..2023
FT                  /*tag= f
FT TATA_signal      2094..2100
FT                  /*tag= g
FT                  /note= "TATA-like sequence; no evidence that this
FT                          functions a promoter"
FT polyA_signal      2514..2520
FT                  /*tag= h
FT misc_signal       2529..2536
FT                  /*tag= i
FT                  /note= "consensus sequence immediately precedes
FT                          polyA-tail (Benoist et al)"

```

XX		
PN	US5482851-A.	
XX		
PD	09-JAN-1996.	
XX		
PF	05-NOV-1993;	93US-0147364.
XX		
PR	13-MAR-1987;	87US-0025423.
PR	22-MAR-1985;	85US-0715142.
PR	04-AUG-1989;	89US-0389929.
PR	04-MAR-1992;	92US-0845893.
PR	05-NOV-1993;	93US-0147364.

XX
PA (GETH) GENENTECH INC.

XX
PI Derynck RMA, Goeddel DV;

XX
DR WPI; 1996-076891/08.
DR P-PSDB; AAR90827.

XX
PT New recombinant human transforming growth factor-beta prods. - produced
PT using Chinese hamster ovary cells, for use in diagnostic applications
PT or in therapy

XX
PS Example 3; Fig 1; 26pp; English.

XX
CC The cDNA encodes the pre-transforming growth factor (TGF) beta 1 protein.
CC The nucleotide sequence was obt'd. by an analysis of several overlapping
CC cDNAs and gene fragments. The DNA is useful for the recombinant
CC production of TGF beta 1, which can be used in, e.g. wound healing.
CC (Revised entry submitted to correct sequence analysis breakdown.)
CC (Updated on 25-MAR-2003 to correct PF field.)

Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 73.5%; Score 994; DB 17; Length 2537;
Best Local Similarity 85.7%; Pred. No. 5.8e-194;
Matches 1148; Conservative 0; Mismatches 145; Indels 47; Gaps 2;

Qy 10 CCGAGATGGGCGCCTTCGGGGCTGCGGCTCTTGCCGTGCTGCTGCCGTGCTGTGGCTGC 69
||| ||| ||| | | | | | | | | | | | | | | | | | | |
Db 837 CCCCCATGCCGCCCTCCGGGGCTGCGGCTGCTGCCGTGCTGCTACCGCTGCTGTGGCTAC 896

Qy 70 TAGTGTCTGACGCCTGGCCGGCCGGCCGGCACTGTCCACTGCAAGACCATCGACATGG 129

Db	897	<div> </div> TGGTGCTGACGCCTGGCCCCGCCGCCGCCGGGACTATCCACCTGCAAGACTATCGACATGG	956
Qy	130	<div> </div> AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	189
Db	957	<div> </div> AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
Qy	190	<div> </div> TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCGCCGCCGCTGCCTGAGGCAGTACTGG	249
Db	1017	<div> </div> TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCGCCGCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	250	<div> </div> CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	309
Db	1077	<div> </div> CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	310	<div> </div> CAGAGGCGGACTACTACGCCAAGGAGGTACCCCGCTGCTAATGGTGGAAAGCGGCAACC	369
Db	1137	<div> </div> CTGAGGCCGACTACTACGCCAAGGAGGTACCCCGCTGCTAATGGTGGAAACCCACAACG	1196
Qy	370	<div> </div> AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	1197	<div> </div> AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	430	<div> </div> AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	1257	<div> </div> AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	1316
Qy	490	<div> </div> GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	549
Db	1317	<div> </div> GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT	1376
Qy	550	<div> </div> GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	609
Db	1377	<div> </div> GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	610	<div> </div> ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	669
Db	1437	<div> </div> ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	670	<div> </div> TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1497	<div> </div> TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	730	<div> </div> TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTG	789
Db	1557	<div> </div> TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGCTG	1616
Qy	790	<div> </div> TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1617	<div> </div> TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	850	<div> </div> CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1676	<div> </div> -----GCCCTGGACACCAACTATTGCT	1697
Qy	910	<div> </div> TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969

Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT	1817
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1818	GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877
Qy	1090	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149
Db	1878	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1937
Qy	1150	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1938	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC	1269
Db	1998	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCCGCCCCGCCCCGGCAGGCCCCGGCCCCAC	2057
Qy	1270	CCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1321
Db	2058	CCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG	2117
Qy	1322	CCCACCTTGGGATCGATTAAA	1341
Db	2118	CCCACCTGGGGCCCCATTAA	2137

RESULT 7

AAQ56923

XX

XX

DT 09-JUL-1994 (first entry)

XX

XX

KW transforming growth factor beta-3; recombinant; wound healing;
KW vulnerary; ss.

XX

XX

FT misc_structure 47..113

FT /*tag= a

```
FT      /note= "possible hairpin loop region"
```

FT CDS 842..2014

FT /*tag= b

FT	mat_peptide	1676..2011
----	-------------	------------

```
FT          /*tag=  c
```

FT	polyA_signal	2515..2521
----	--------------	------------

FT /*tag= d
XX
PN US5284763-A.
XX
PD 08-FEB-1994.
XX
PF 04-MAR-1992; 92US-0845893.
XX
PR 22-MAR-1985; 85US-0715142.
PR 13-MAR-1987; 87US-0025423.
PR 04-AUG-1989; 89US-0389929.
PR 04-MAR-1992; 92US-0845893.
XX

PA (GETH) GENENTECH INC.
XX

PI Derynk RMA, Goeddel DV;
XX

DR WPI; 1994-056343/07.
DR P-PSDB; AAR46227.
XX

PT Nucleic acid sequences encoding transforming growth factor-beta -
PT diagnostic probes, and for use in therapeutics
XX

PS Disclosure; Fig 1b; 25pp; English.
XX

CC cDNA sequences were determined for human pre-TGF-beta-1 (AAQ56923),
CC pig TGF-beta-3 (AAQ56925) and human TGF-beta-3 (AAQ56926), and the
CC corresponding amino acid sequences were determined (AAR46227-29,
CC respectively). A genomic fragment corresponding to a human TGF-
CC beta-1 exon (AAQ56924) was also isolated and its amino acid sequence
CC determined (AAR46230). The sequences have been used in the
CC construction of vectors for the expression of recombinant TGF-
CC beta.
XX

CC (Updated on 25-MAR-2003 to correct PF field.)
XX

SQ Sequence 2537 BP; 473 A; 890 C; 742 G; 432 T; 0 other;

Query Match 73.3%; Score 992.4; DB 15; Length 2537;
Best Local Similarity 85.6%; Pred. No. 1.2e-193;
Matches 1147; Conservative 0; Mismatches 146; Indels 47; Gaps 2;

Qy 10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
|| ||| |||| || ||||||||| ||||||||| ||||||||| ||
Db 837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy 70 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 129
| ||||||||||||||| ||||||||| ||||||||| ||||||||| ||
Db 897 TGGTGCTGACGCCTGGCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy 130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 189
||||||||||||||||| ||||||||| ||||||||| ||||||||| ||
Db 957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy 190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 249
| ||||||||||||||| ||||||||| ||||||||| ||||||||| ||
Db 1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 1076

Qy 250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGC 309
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACGGGAGCCCGAGC 1136

 Qy 310 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC 369
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1137 CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACG 1196

 Qy 370 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

 Qy 430 AGCTCCGGGAAGCGGTGCGCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 1316

 Qy 490 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT 549
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1317 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT 1376

 Qy 550 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG 609
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1377 GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG 1436

 Qy 610 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC 669
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1437 ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC 1496

 Qy 670 TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 729
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1497 TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT 1556

 Qy 730 TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGTC 789
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1557 TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGTC 1616

 Qy 790 TCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG 849
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1617 TTCTCATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA- 1675

 Qy 850 CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT 909
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1676 -----GCCCTGGACACCAACTATTGCT 1697

 Qy 910 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC 969
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1698 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC 1757

 Qy 970 TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1029
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1758 TCGGCTGGAAGTGGATCCACGAGCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCT 1817

 Qy 1030 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1089
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1818 GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1877

 Qy 1090 ACAACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA 1149

Db	1878	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCG CAGGCGCTGGAGCCGCTGCCCA	1937
Qy	1150	TCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1938	TCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCGCCCGCCACAGCCCGCCCAACCCGGCAGGCCCGGC	1269
Db	1998	CCTGCAAGTGCAGCTGAGGTCCCGCCCGCCCGCCCGCCCGCCCGGCAGGCCCGGCCCCAC	2057
Qy	1270	CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1321
Db	2058	CCCGCCCGCCCGCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG	2117
Qy	1322	CCCACTTGGGATCGATTAAA	1341
Db	2118	CCCACCTGGGGCCCCATTAA	2137

RESULT 8

AAV52933

ID AAV52933 standard; cDNA; 2537 BP.

XX

AC AAV52933;

XX

DT 25-MAR-2003 (updated)

DT 21-DEC-1998 (first entry)

XX

DE Human pre-transforming growth factor-beta 1 cDNA.

XX

KW Transforming growth factor-beta 1; TGF-beta 1; human; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT CDS 842..2014

FT /*tag= a

FT mat peptide 1676..2011

FT /*tag= b

```
FT      stem_loop      37..113
```

FT /*tag= b

```
FT      /note= "putative stable hairpin loop"
```

```
FT    misc_feature    2015..2100
```

FT /*tag= c

```
FT                               /note= "GC-rich sequence"
```

FT	polyA_signal	2514..2520
----	--------------	------------

```
FT                               /*tag=  d
```

XX

PN US5801231-A.

XX

PD 01-SEP-1998.

XX

PF 30-MAY-1995; 95US-0454468.

XX

PR 13-MAR-1987; 87US-0025423.

PR 22-MAR-1985; 85US-0715142.

Qy	250	CTCTTTTACAACAGTATCCCGCGACCCGGGTAGCCGGGGAAAGTGTGCGAACCCGGAGCCCGGAGC	309
Db	1077	CCCTGTACAACAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGCAGAACCCGGAGCCCGGAGC	1136
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	369
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	370	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT	549
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT	1376
Qy	550	GGCGCTACCTCAGCAACCCGGCTGCTGGCCCCCAGTGA CTACCCGGAGTGGCTGTCCTTTG	609
Db	1377	GGCGATACCTCAGCAACCCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	610	ATGTCACCCGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	669
Db	1437	ATGTCACCCGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGAGAAATTGAGGGCTTTTCGCC	1496
Qy	670	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	789
Db	1557	TCACTACCCGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTCTGC	1616
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	970	TGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCT	1817
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1818	GCCCCATACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877
Qy	1090	ACAACCCGGGCGCGTGGCGGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149

Db	1878	ATAAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCG	1937
Qy	1150	TCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1938	TCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCCAACCGGCAGGCCCCGGC	1269
Db	1998	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGGCAGGCCCCGGCCCCAC	2057
Qy	1270	CCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1321
Db	2058	CCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG	2117
Qy	1322	CCCACCTGGGATCGATTAAA	1341
Db	2118	CCCACCTGGGGCCCCATTAA	2137

AAQ03268

XX

XX

DT 12-AUG-1990 (first entry)

XX

XX

XX

XX

FT	sig peptide	283..324
----	-------------	----------

```
FT          /*tag=  a
```

FT	mat_peptide	1096..1431
----	-------------	------------

```
FT          /*tag=  b
```

FT /product=human transforming growth factor-beta

XX

XX

XX

XX

XX

XX

XX

DR P-PSDB; AAR03743.

XX

PT Inhibition of proliferation of epidermal cells -
PT used to treat psoriasis by contacting cells with compositions
PT containing transforming growth factor-beta.

XX

PS Disclosure; fig 1; 20pp; English.

XX

CC TGF-beta may be used in the treatment of hyperplasia
CC associated with acanthosis-categorised skin diseases, and
CC in alleviating psoriatic symptoms associated with cytokine-
CC induced phenomena. See also AAQ03269 and AAR03750.
CC (Updated on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 1561 BP; 301 A; 547 C; 446 G; 267 T; 0 other;

Query Match 72.9%; Score 986.2; DB 11; Length 1561;
Best Local Similarity 85.8%; Pred. No. 2.1e-192;
Matches 1150; Conservative 0; Mismatches 143; Indels 48; Gaps 3;

```
Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      257 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCGGCTGCTGCTACCGCTGCTGTGGCTAC 316

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 129
      |  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      317 TGGTGCTGACGCCTAGCCGGCCGGCCGAGGACTATCCACCTGCAAGACTATCGACATGG 376

Qy     130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     377 AGCTGGTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 436

Qy     190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 249
      |  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     437 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 496

Qy     250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCCGAGC 309
      |  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     497 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCGGAGCCCGAAC 556

Qy     310 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACC 369
      |  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     557 CGGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACG 616

Qy     370 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     617 AAATCTATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAG 676

Qy     430 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     677 AGCTCCGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 736

Qy     490 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     737 GGCTCAAGTTAAAAGTGGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTT 796

Qy     550 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG 609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     797 GGCGATACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGAGTGGTTGTCTTTTG 856
```


XX
 AC AAI58342;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 545.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM39186.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 545; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX

SQ Sequence 2742 BP; 526 A; 938 C; 799 G; 479 T; 0 other;

Query Match 72.8%; Score 984.6; DB 22; Length 2742;
Best Local Similarity 85.7%; Pred. No. 4.9e-192;
Matches 1149; Conservative 0; Mismatches 144; Indels 48; Gaps 3;

```
Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGCTGCTGTGGCTGC 69
      ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     837 CCCCCATGCCGCCCTCCGGGGCTGCGGCTGCTGCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     897 TGGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy     130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy     190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 249
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCTGTCTCG 1076

Qy     250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGGAGCCCGAGC 309
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAAACCGGAGCCCGAGC 1136

Qy     310 CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC 369
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1137 CTGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACG 1196

Qy     370 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy     430 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1257 AGCTCCGAGAAGCGGTACCTGAAACCGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 1316

Qy     490 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT 549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1317 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT 1376

Qy     550 GGCGTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGCGTGTCTCTTTG 609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1377 GGCGTACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTTATCTTTTG 1436

Qy     610 ATGTCACCGGAGTTGTGCGGCAGTGCGTGACCCGCGAGAGGGCTATAGAGGGTTTTTCGCC 669
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1437 ATGTCACCGGAGTTGTGCGGCAGTGCGTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC 1496

Qy     670 TCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 729
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

XX
KW Hypertension therapy; hypotensive agent; blood pressure modulator;
KW ss.
XX
OS Monkey.
XX
FH Key Location/Qualifiers
FT CDS 262..282
FT /*tag= a
FT sig_peptide 283..324
FT /*tag= b
FT CDS 325..1098
FT /*tag= c
FT mat_peptide 1099..1436
FT /*tag= d
XX
PN WO9119513-A.
XX
PD 26-DEC-1991.
XX
PF 20-JUN-1991; 91WO-US04449.
XX
PR 20-JUN-1990; 90US-0541221.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Oleson FB, Comereski CR;
XX
DR WPI; 1992-024199/03.
DR P-PSDB; AAR20124.
XX
PT Use of transforming growth factor (TGF)-beta and their
PT antagonists - for modulating blood pressure, for treating
PT hypertension and hypotension
XX
PS Disclosure; Fig 1; 42pp; English.
XX
CC A new method for treating hypertension comprises administering a
CC transforming growth factor (TGF)-beta to an individual at a dose
CC effective for lowering blood pressure; the TGF-beta may be e.g.
CC mature TGF-beta, TGF-beta2, a mature TGF-betal/beta2 hybrid, TGF-
CC betal precursor, a latent TGF-beta2 precursor, hybrid TGF-betal/TGF-
CC beta2 precursor, a latent TGF-betal complex or a latent TGF-beta2
CC complex.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1559 BP; 300 A; 546 C; 446 G; 267 T; 0 other;

Query Match 72.7%; Score 983.8; DB 13; Length 1559;
Best Local Similarity 85.8%; Pred. No. 6.6e-192;
Matches 1146; Conservative 0; Mismatches 142; Indels 47; Gaps 3;

Qy 15 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 74
||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320

Qy 75 CTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 134

```

Db	321	 CTGACGCCTAGCCGGCCGGCCGAGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	380
Qy	135	GTGAAGCGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	194
Db	381	 GTGAAGCGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	440
Qy	195	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTT	254
Db	441	 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	255	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG	314
Db	501	 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCCGAGCCCGAACCGGAG	560
Qy	315	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	374
Db	561	 GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	375	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	434
Db	621	 TATGACAAGTTCAGGAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	435	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	494
Db	681	 CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	495	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACAGAAATACAGCAATGATTCCTGGCGC	554
Db	741	 AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	800
Qy	555	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGAAGTCAACCGGAGTGGCTGTCTTTGATGTC	614
Db	801	 TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	615	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCTCAGT	674
Db	861	 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCTTAGC	920
Qy	675	GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	734
Db	921	 GCCCACTGTTCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	735	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	794
Db	981	 ACCGGCCGCCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCCTGCTTCTC	1040
Qy	795	ATGGCCACCCCGCTGGAGAGGGCCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	854
Db	1041	 ATGGCCACCCCGCTGGAGAGGGCCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-----	1094
Qy	855	GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC	914
Db	1095	----- GCCCTGGACACCAACTACTGCTTCAGC	1121
Qy	915	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC	974

Db	1122	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGC	1181
Qy	975	TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1034
Db	1182	TGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTGGGGCCCTGTCCC	1241
Qy	1035	TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAAC	1094
Db	1242	TACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAAC	1301
Qy	1095	CCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1154
Db	1302	CCGGGCGCCTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1361
Qy	1155	TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC	1214
Db	1362	TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC	1421
Qy	1215	AAGTGCAGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCCGGCAGGCCCCGCCCCAC	1274
Db	1422	AAATGCAGCTGAGGCCCCGCCCCGCCCCGCCCCACCCCGGCAGGCCCCGCCCCACCC	1481
Qy	1275	CCCCGCCCCG-----CTCACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAAGCCCCAC	1326
Db	1482	CCACCCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAAGCCCCAC	1541
Qy	1327	TTGGGATCGATTAAA	1341
Db	1542	CTGGGGCCCCATTAA	1556

RESULT 12

AAQ13392

ID AAQ13392 standard; DNA; 1821 BP.

XX

AC AAQ13392;

XX

DT 20-NOV-1991 (first entry)

XX

DE Human pro-TGF-beta 1 gene.

XX

KW Osteogenetic; tumoricidal; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	512..1684
----	-----	-----------

FT		/*tag= a
----	--	----------

FT	sig_peptide	512..598
----	-------------	----------

FT		/*tag= b
----	--	----------

FT	misc_RNA	599..1684
----	----------	-----------

FT		/*tag= c
----	--	----------

FT		/note= "pro-TGF-beta 1"
----	--	-------------------------

FT	mat_peptide	1346..1684
----	-------------	------------

FT		/*tag= e
----	--	----------

FT		/note= "TGF-beta 1"
----	--	---------------------

XX

Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGTCTGA	489
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT	549
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT	1046
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCCTTTG	609
Db	1047	GGCGATACTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTG	1106
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	669
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	670	TCAGTGCCCACTGTTCTGTGTACAGCAAAGATAACACACTCCAGTGGAATTAACGGGT	729
Db	1167	TTAGCGCCCACTGCTCCTGTGTACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1226
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	789
Db	1227	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1286
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1346	-----GCCCTGGACACCAACTATTGCT	1367
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1368	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1427
Qy	970	TGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1428	TGGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1487
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1488	GCCCCACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1547
Qy	1090	ACAACCCGGGCGCGTGGCGGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149
Db	1548	ATAACCCGGGCGCCTGGCGGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1607
Qy	1150	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1608	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1667
Qy	1210	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC	1269
Db	1668	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCCGCCCGCCCCGGCAGGCCCGGCCAC	1727
Qy	1270	CCCACCCCCGCGCGCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1329

```

      ||| |||| ||| ||      || ||||| ||||| ||||| ||||| ||||| ||
Db      1728 CCCGCCCCGCCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCCAA 1787

Qy      1322 CCC 1324
      ||
Db      1788 GCC 1790

```

RESULT 13

AAN81084

ID AAN81084 standard; cDNA; 1560 BP.

XX

AC AAN81084;

XX

DT 25-MAR-2003 (updated)

DT 09-OCT-1990 (first entry)

XX

DE Coding sequence of simian transforming growth factor-beta 1.

XX

KW Transforming growth factor-beta 1; tumour treatment; ss cDNA.

XX

OS Cercopithecus aethiops.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	261..1433
----	-----	-----------

FT		/*tag= a
----	--	----------

FT	sig_peptide	282..323
----	-------------	----------

FT		/*tag= b
----	--	----------

FT	mat_peptide	1095..1433
----	-------------	------------

FT		/*tag= c
----	--	----------

XX

PN EP293785-A.

XX

PD 07-DEC-1988.

XX

PF 27-MAY-1988; 88EP-0108528.

XX

PR 29-MAY-1987; 87US-0055662.

PR 25-JAN-1988; 88US-0147842.

XX

PA (ONCO) ONCOGEN.

PA (BRIM) BRISTOL-MYERS CO.

XX

PI Purchio AG, Gentry L, Twardzik D;

XX

DR WPI; 1988-347488/49.

DR P-PSDB; AAP80647.

XX

PT Prodn. of simian transforming growth factor beta-1 - by culturing
PT transfected eucaryotic cells, and new precursor proteins, useful for
PT treating tumours.

XX

PS Disclosure; Page ?; pp; English.

XX

CC The cDNA is prepd. from African green monkey cell line BSC-40 and is
CC expressed in eukaryotic cells in plasmid pSV2. There is 100% homology
CC between mature simian and human TGF-beta 1. The plasmid also contains

SQ Sequence 1560 BP; 301 A; 547 C; 445 G; 267 T; 0 other;

Qy	15	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	74
Db	261	ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG	320
Qy	75	CTGACGCCTTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	134
Db	321	CTGACGCCTAGCCGGCCGGCCGAGGACTATCCACCTGCAAGACTATCGACATCGAGCTG	380
Qy	135	GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	194
Db	381	GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	440
Qy	195	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	254
Db	441	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	255	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG	314
Db	501	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCGGAGCCCGAACCGGAG	560
Qy	315	GCGGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATC	374
Db	561	GCCGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAACCCACAACGAAATC	620
Qy	375	TATGATAAAATTCAAGGGCACCCCCCAAGCTTATATATGCTGTTCAACACGTGCGAGCTC	434
Db	621	TATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	435	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	494
Db	681	CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	495	AAGTTAAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	554
Db	741	AAGTTAAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	800
Qy	555	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGATGTC	614
Db	801	TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	615	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT	674
Db	861	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	920
Qy	675	GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	734
Db	921	GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980

Qy	735	TCTGGCCGCCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCCTGCTCCTC	794
Db	981	ACCGGCCGCCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCCTGCTTCTC	1040
Qy	795	ATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGTCCCGGCACCGCCGAGCCCTG	854
Db	1041	ATGGCCACCCCGCTGGAGAGGGGCCAACATCTGCAAAGTCCCGGCACCGCCGA-----	1094
Qy	855	GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC	914
Db	1095	-----GCCCTGGACACCAACTACTGCTTCAGC	1121
Qy	915	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC	974
Db	1122	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGC	1181
Qy	975	TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1034
Db	1182	TGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1241
Qy	1035	TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGGCTCTGTACAACCAGCACAAAC	1094
Db	1242	TACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAAC	1301
Qy	1095	CCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1154
Db	1302	CCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1361
Qy	1155	TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC	1214
Db	1362	TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGA	1421
Qy	1215	AAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGCCCCAC	1274
Db	1422	AAATGCAGCTGAGGCCCCGCCCCGCCACCCCGGCCAGGCCCCGGCCCCGCCCCAC	1481
Qy	1275	CCCCGCCCCCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAAGCCCA	1325
Db	1482	CCCACCCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAAGCCCA	1541
Qy	1326	CTTGGGATCGATTAAA	1341
Db	1542	CCTGGGGCCCCATTAA	1557

XX
 KW HIV; AIDS; SIV; vaccine; AZT; CD4; cytokines; growth
 KW factors; ds.
 XX
 OS Cebus apella.
 XX
 FH Key Location/Qualifiers
 FT CDS 267..1437
 FT /*tag= a
 FT mat_peptide 1103..1437
 FT /*tag= b
 XX
 PN EP356935-A.
 XX
 PD 07-MAR-1990.
 XX
 PF 25-AUG-1989; 89EP-0115719.
 XX
 PR 25-AUG-1988; 88US-0236698.
 XX
 PA (ONCO) ONCOGEN LP.
 XX
 PI Brankovan V, Lioubin M, Purchio A;
 XX
 DR WPI; 1990-068723/10.
 DR P-PSDB; AAR05663.
 XX
 PT Compsns. contg. transforming growth factor beta -
 PT used for inhibitions of HIV infection and replication in vivo.
 XX
 PS Disclosure; Fig 1; 20pp; English.
 XX
 CC TGF-beta may be used in vivo to prevent formation of syncytia and
 CC inhibit HIV infection. TGF may also be used with other HIV treatments
 CC (AZT, soluble CD4 etc.).
 CC (Updated on 09-JAN-2003 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1560 BP; 301 A; 547 C; 445 G; 267 T; 0 other;

Query Match 72.6%; Score 982.8; DB 11; Length 1560;
 Best Local Similarity 85.8%; Pred. No. 1.1e-191;
 Matches 1146; Conservative 0; Mismatches 142; Indels 48; Gaps 3;

Qy 15 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 74
 |||||
 Db 261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320
 Qy 75 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 134
 |||||
 Db 321 CTGACGCCTAGCCGGCCGGCCCGGAGACTATCCACCTGCAAGACTATCGACATCGAGCTG 380
 Qy 135 GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 194
 |||||
 Db 381 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 440
 Qy 195 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 254

Db	441	 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	255	TACAACAGTACCCCGGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGCCAGAG	314
Db	501	 TACAACAGCACCCCGGACCGGGTGGCCGGGGAGAGTGCAGGAGCCCGAGCCCGAACCGGAG	560
Qy	315	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	374
Db	561	 GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	375	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	434
Db	621	 TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	435	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	494
Db	681	 CGAGAAGCAGTACCTGAACCTGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	495	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	554
Db	741	 AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	800
Qy	555	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	614
Db	801	 TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	615	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	674
Db	861	 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	920
Qy	675	GCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	734
Db	921	 GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	735	TCTGGCCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	794
Db	981	 ACCGGCCCGCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCTGCTTCTC	1040
Qy	795	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	854
Db	1041	 ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-----	1094
Qy	855	GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC	914
Db	1095	----- -----GCCCTGGACACCAACTACTGCTTCAGC	1121
Qy	915	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC	974
Db	1122	 TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGC	1181
Qy	975	TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1034
Db	1182	 TGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1241
Qy	1035	TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAAC	1094

```

Db      1242 TACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAAC 1301
Qy      1095 CCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG 1154
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1302 CCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG 1361
Qy      1155 TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC 1214
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1362 TACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGA 1421
Qy      1215 AAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGCCCCAC 1274
          ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1422 AAATGCAGCTGAGGCCCCGCCCCGCCACCCCGGCAGGCCCCGGCCCCGCCAC 1481
Qy      1275 CCCC GCCCGCCT-----CACCGGGCTGTATTTAAGGACA-TCGTGCCCCAAGCCCA 1325
          |||  |||  ||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1482 CCCACCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAAGCCCA 1541
Qy      1326 CTTGGGATCGATTAAA 1341
          |||||  ||  ||
Db      1542 CCTGGGGCCCCATTAA 1557

```

RESULT 15

AAT05876

ID AAT05876 standard; cDNA; 2745 BP.

XX

AC AAT05876;

XX

DT 25-JUN-1996 (first entry)

XX

DE cDNA encoding transforming growth factor-beta 1.

XX

KW macrophage inducible nitric oxide synthase; iNOS; constitutive NOS;

KW interleukin-1-beta; transforming growth factor-beta; TGF-beta; IL1-beta;

KW nitric oxide production; hypotension; inflammation; septic shock;

KW treatment; ds.

XX

OS Mammalian sp.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	842..2017
----	-----	-----------

FT		/*tag= a
----	--	----------

FT		/product= transforming growth factor-beta 1
----	--	---

XX

PN WO9526745-A1.

XX

PD 12-OCT-1995.

XX

PF 05-APR-1994; 94WO-US03705.

XX

PR 05-APR-1994; 94WO-US03705.

XX

PA (HARD) HARVARD COLLEGE.

XX

PI Lee M, Perrella MA;

XX

DR WPI; 1995-358443/46.
 DR P-PSDB; AAR83054.
 XX
 PT Treatment of hypotension, esp. in septic shock - by administering
 PT transforming growth factor-beta e.g. to inhibit inducible nitric
 PT oxide synthase gene transcription
 XX
 PS Disclosure; Fig 15; 52pp; English.
 XX
 CC The cDNA encodes transforming growth factor-beta 1 (TGF-beta 1) which
 CC has been found to inhibit inducible nitric oxide synthase (iNOS) gene
 CC transcription, esp. in interleukin-1-beta (IL1-beta) stimulated rat
 CC smooth muscle cells, and at a dose which does not inhibit constitutive
 CC NOS. TGF-beta 1 or 2 (AAR83055) or their active fragments, can be used
 CC in the treatment of hypotension, such as that associated with severe
 CC inflammation or septic shock.
 XX
 SQ Sequence 2745 BP; 527 A; 938 C; 801 G; 479 T; 0 other;

Query Match 72.5%; Score 981; DB 16; Length 2745;
 Best Local Similarity 85.5%; Pred. No. 2.7e-191;
 Matches 1148; Conservative 0; Mismatches 145; Indels 50; Gaps 3;

```

Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      897 TGGTGCTGACGCCTGGCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy     130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy     190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 249
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCAGGCCGTGCTCG 1076

Qy     250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 309
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy     310 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAAGCGGCAACC 369
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1137 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAACCCACAACG 1196

Qy     370 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy     430 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT-- 487
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 1316

Qy     488 -GAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATT 546
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 1317 GGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATT 1376

Qy 547 CCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCCT 606
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1377 CCTGGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTT 1436

Qy 607 TTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTC 666
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1437 TTGATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTC 1496

Qy 667 GCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACG 726
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1497 GCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACG 1556

Qy 727 GGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC 786
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1557 GGTTCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCC 1616

Qy 787 TGCTCCTCATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCC 846
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1617 TGCTTCTCATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCC 1676

Qy 847 GAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACT 906
 || ||||| ||||| ||||| |||||

Db 1677 GA-----GCCCTGGACACCAACTATT 1697

Qy 907 GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGG 966
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1698 GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGG 1757

Qy 967 ACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGC 1026
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1758 ACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGC 1817

Qy 1027 CCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGTGCTGTGTAACAC 1086
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1818 CCTGCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGTGCTGTGTAACAC 1877

Qy 1087 AGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGC 1146
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1878 AGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGC 1937

Qy 1147 CCATCGTGTACTACGTGGGCGCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC 1206
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1938 CCATCGTGTACTACGTGGGCGCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC 1997

Qy 1207 GTTCCTGCAAGTGACAGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCGGCAGGCC 1266
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1998 GTCCTGCAAGTGACAGCTGAGGTCCCGCCCCGCCCCGCCCCGCCCCGCGCAGGCCCGGCC 2057

Qy 1267 GGCCCCACCCCCGCCCCCT-----CACCGGGCTGTATTTAAGGACATCGTGCCCC 1318
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2058 CACCCGCCCCGCCCCGCTGCTTGGCCATGGGGCTGTATTTAAGGACACCGTGCCCC 2117

Qy 1319 AAGCCCACTTGGGATCGATTAAA 1341
 ||||| ||||| ||||| ||||| |||||

Db 2118 AAGCCCACTTGGGGCCCCATTAA 2140

Search completed: October 27, 2003, 19:11:23
Job time : 416.484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 18:35:27 ; Search time 402.152 Seconds
(without alignments)
9022.658 Million cell updates/sec

Title: US-10-017-372E-34
Perfect score: 1353
Sequence: 1 gatctggtaccgagatggcg.....cgattaaagcggccgcgact 1353

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%	Query		DB	ID	Description
No.	Score	Match	Length					
	1	985	72.8	1821	14	US-10-087-268-1		Sequence 1, Appli
	2	984.6	72.8	2742	14	US-10-037-270-220		Sequence 220, App
	3	983.4	72.7	1821	14	US-10-087-268-4		Sequence 4, Appli
	4	981	72.5	2745	11	US-09-948-002-28		Sequence 28, Appl
	5	862.2	63.7	2094	11	US-09-948-002-1		Sequence 1, Appli
	6	842.6	62.3	1585	11	US-09-948-002-27		Sequence 27, Appl
	7	657.2	48.6	1376	10	US-09-756-283A-19		Sequence 19, Appl
	8	594.6	43.9	1352	10	US-09-756-283A-21		Sequence 21, Appl
	9	366.4	27.1	489	11	US-09-911-904-167		Sequence 167, App
	10	308.6	22.8	339	10	US-09-813-271B-1		Sequence 1, Appli
	11	241.6	17.9	2574	11	US-09-906-158-3		Sequence 3, Appli
	12	241.6	17.9	2574	13	US-10-028-158-20		Sequence 20, Appl
	13	236.4	17.5	4382	12	US-09-957-458B-9		Sequence 9, Appli
	14	233.2	17.2	2879	11	US-09-906-158-10		Sequence 10, Appl
	15	227.2	16.8	336	10	US-09-813-271B-7		Sequence 7, Appli
	16	209.4	15.5	339	10	US-09-813-271B-5		Sequence 5, Appli
	17	188	13.9	4267	11	US-09-948-002-47		Sequence 47, Appl
c	18	185.4	13.7	2381	12	US-10-311-455-2384		Sequence 2384, Ap
	19	182.4	13.5	336	10	US-09-813-271B-11		Sequence 11, Appl
	20	181.2	13.4	2381	12	US-10-311-455-2383		Sequence 2383, Ap
	21	180.8	13.4	336	10	US-09-813-271B-9		Sequence 9, Appli
	22	160.4	11.9	2570	12	US-09-960-706-663		Sequence 663, App
	23	160.4	11.9	2912	13	US-10-044-090-323		Sequence 323, App
	24	158.2	11.7	339	10	US-09-813-271B-3		Sequence 3, Appli
c	25	141.6	10.5	597	9	US-09-864-761-15319		Sequence 15319, A
c	26	134.8	10.0	154	9	US-09-864-761-31841		Sequence 31841, A
c	27	131.8	9.7	206	10	US-09-833-381-577		Sequence 577, App
c	28	123.2	9.1	537	9	US-09-864-761-8844		Sequence 8844, Ap
c	29	122.2	9.0	148	9	US-09-864-761-25510		Sequence 25510, A
	30	117	8.6	181	10	US-09-833-381-1416		Sequence 1416, Ap
	31	104.8	7.7	851	13	US-10-027-632-152938		Sequence 152938,
	32	102.2	7.6	29000	11	US-09-906-158-17		Sequence 17, Appl
c	33	100.4	7.4	224	10	US-09-833-381-73		Sequence 73, Appl
	34	79.4	5.9	450	11	US-09-918-995-6145		Sequence 6145, Ap
c	35	69.8	5.2	431	9	US-09-864-761-18644		Sequence 18644, A
c	36	69.4	5.1	176	10	US-09-833-381-571		Sequence 571, App
	37	69	5.1	584	12	US-10-029-386-9758		Sequence 9758, Ap
	38	68.2	5.0	176	12	US-10-029-386-23458		Sequence 23458, A
c	39	64	4.7	363	10	US-09-833-381-585		Sequence 585, App
	40	62.6	4.6	364	12	US-10-029-386-26476		Sequence 26476, A
	41	62.6	4.6	544	12	US-10-029-386-12776		Sequence 12776, A
	42	62.6	4.6	927	12	US-10-244-718-1		Sequence 1, Appli
	43	57.4	4.2	486	11	US-09-918-995-25641		Sequence 25641, A
	44	55.2	4.1	594	13	US-10-027-632-141376		Sequence 141376,
	45	53.2	3.9	658	9	US-09-765-527-254		Sequence 254, App

ALIGNMENTS

RESULT 1

US-10-087-268-1

; Sequence 1, Application US/10087268

```

; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsonn, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or
condition
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(511)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (512)..(1684)
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (512)..(598)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (1685)..(1821)
; OTHER INFORMATION:
US-10-087-268-1

```

```

Query Match          72.8%; Score 985; DB 14; Length 1821;
Best Local Similarity 85.9%; Pred. No. 5.9e-255;
Matches 1136; Conservative 0; Mismatches 140; Indels 47; Gaps 2;

```

```

Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      || ||| |||| || ||||||||||| ||||||||||| ||||||||||| ||
Db      507 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 566

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||||||||||||||||||||||| ||||||||||| ||||||||||| ||
Db      567 TGGTGCTGACGCCTGGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 626

Qy      130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||||||||||||||||||||||| ||||||||||| ||||||||||| ||||
Db      627 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 686

Qy      190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 249
      | ||||||||||||||||||||||| ||||||||||| ||||||||||| ||||
Db      687 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 746

Qy      250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 309
      | || ||||||||| ||||||||||| ||||||||| |||| ||||||||||| ||
Db      747 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 806

Qy      310 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 369
      | ||||| ||||||||||| ||||||||||| ||||||||||| ||||| ||||
Db      807 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG 866

```


Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	867	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	926
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	549
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1046
Qy	550	GGCGCTACCTCAGCAACCGGTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	609
Db	1047	GGCGATACCTCAGCAACCGGTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1106
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	669
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	670	TCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1167	TTAGCGCCCACTGCTCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1226
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGTC	789
Db	1227	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTTCATGGCATGAACCGGCCCTTCTGTC	1286
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1346	-----GCCCTGGACACCAACTATTGCT	1367
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1368	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1427
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1428	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCT	1487
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1488	GCCCCACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1547
Qy	1090	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	1149
Db	1548	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCA	1607
Qy	1150	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209
Db	1608	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1667

```
Qy      1210 CCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCGCCCACCCGGCAGGCCCGGC 1269
         |||||||
Db      1668 CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCCGCCCGGCAGGCCCGGCCAAC 1727

Qy      1270 CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG 1321
         |||||
Db      1728 CCCGCCCCGCCCGCGCTGCGCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAA 1787

Qy      1322 CCC 1324
         ||
Db      1788 GCC 1790
```

US-10-037-270-220

; Publication No. US20030104529A1

; APPLICANT: Tang, Y. Tom

APPLICANT: Asundi, Vinod

; APPLICANT: Ren, Feiyan

APPLICANT: Zhao, Qing A.

; APPLICANT: Xue, Aidong

APPLICANT: Wang, Jian-Rui

; APPLICANT: Ma, Yunging

; APPLICANT: Wang, Zhiwei

APPLICANT: Drmanac, Radoje T.

10 TITLE OF INVENTION: Polypeptides

CURRENT APPLICATION NUMBER: US/10/037,270

; PRIOR APPLICATION NUMBER: 09/552,317

PRIOR APPLICATION NUMBER: 09/

```

; NUMBER OF SEQ ID NOS: 1104

```

SEO ID NO 220

TYPE: DNA

FEATURE:

; LOCATION: (842) .. (2014)

US-10-037-270-220

Query Match 72.8%; Score 984.6; DB 14; Length 2742;

Best Local Similarity 85.7%; Pred. No. 8.4e-255;

Matches 1149; Conservative 0; Mismatches 144; Indels 48; Gaps 3;

```
Qy      10 CCGAGATGGCGCCTTCGGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      ||  ||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     897 TGGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTATCCACCTGCAAGACTATCGACATGG 956

Qy     130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy     190 TTGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 249
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1017 TCGCCAGCCCCCGAGCCAGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 1076

Qy     250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGAGCCCGAGC 309
      | || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGAGCCCGAGC 1136

Qy     310 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC 369
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1137 CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACG 1196

Qy     370 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy     430 AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 1316

Qy     490 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT 549
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1317 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT 1376

Qy     550 GCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTTG 609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1377 GCGGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTG 1436

Qy     610 ATGTCACCGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC 669
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1437 ATGTCACCGAGTTGTGCGGCAGTGGTTGAGCCGTGAGGGGAAATTGAGGGCTTTTCGCC 1496

Qy     670 TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 729
      || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1497 TTAGCGCCCACTGCTCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT 1556

Qy     730 TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCTGCTGC 789
      ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1557 TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCCTTCTGCTGC 1616

Qy     790 TCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG 849
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1617 TTCTCATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA- 1675
```

```

QY      850 CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT 909
          ||||| ||||| |||||
Db      1676 -----GCCCTGGACACCAACTATTGCT 1697

QY      910 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC 969
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1698 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC 1757

QY      970 TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1029
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1758 TCGGCTGGAAGTGGATCCACGAGCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCT 1817

QY     1030 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1089
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1818 GCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1877

QY     1090 ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCCA 1149
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1878 ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCGCTGCCCCA 1937

QY     1150 TCGTGTAATACTGCGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1209
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1938 TCGTGTAATACTGCGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT 1997

QY     1210 CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC 1269
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1998 CCTGCAAGTGCAGCTGAGGTCCC GCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC 2057

QY     1270 CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAA 1320
          ||| |||| ||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db     2058 CCCGCCCCGCCCCGCTGCCCTTGCCCATGGGGGCTGTATTTAAGGACACCCGTGCCCCAA 2117

QY     1321 GCCCACTTGGGATCGATTAAA 1341
          ||||| |||| | | ||
Db     2118 GCCCACCTGGGGCCCCATTAA 2138

```

RESULT 3

US-10-087-268-4

; Sequence 4, Application US/10087268

; Publication No. US20030119010A1

; GENERAL INFORMATION:

; APPLICANT: Jonsonn, Julie Ruth

; APPLICANT: Powell, Elizabeth Ellen

; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or condition

; FILE REFERENCE: Fibrosis

; CURRENT APPLICATION NUMBER: US/10/087,268

; CURRENT FILING DATE: 2002-03-01

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1821

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

```
; NAME/KEY: 5'UTR
; LOCATION: (1)..(511)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (512)..(1684)
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (512)..(598)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (1685)..(1821)
; OTHER INFORMATION:
```

US-10-087-268-4

```
Query Match          72.7%; Score 983.4; DB 14; Length 1821;
Best Local Similarity 85.8%; Pred. No. 1.6e-254;
Matches 1135; Conservative 0; Mismatches 141; Indels 47; Gaps 2;
```

```
Qy      10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      || ||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      507 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 566

Qy      70 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 129
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      567 TGGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTATCCACCTGCAAGACTATCGACATGG 626

Qy     130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      627 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATCCTGTCCAAGCTGCGGC 686

Qy     190 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCGAGTACTGG 249
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      687 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCCAGGCGCGTGCTCG 746

Qy     250 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTGAAACCGGAGCCCGAGC 309
      | || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      747 CCCTGTACAACAGCACCCGCGACCGGGTGCCGGGGAGAGTGCAGAAACCGGAGCCCGAGC 806

Qy     310 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGAAAGCGGCAACC 369
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      807 CTGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGAAACCCACAACG 866

Qy     370 AAATCTATGATAAAATTCAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 429
      ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db      867 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 926

Qy     430 AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 489
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      927 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 986

Qy     490 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 549
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      987 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT 1046

Qy     550 GGCGTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG 609
      |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1047 GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTTATCTTTTG 1106
```



```

; Publication No. US20030050265A1
; GENERAL INFORMATION:
;   APPLICANT: Nicholas M. Dean
;   APPLICANT: Susan F. Murray
;   TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
;   TITLE OF INVENTION: FACTOR BETA EXPRESSION
;   FILE REFERENCE: ISPH-0607
;   CURRENT APPLICATION NUMBER: US/09/948,002
;   CURRENT FILING DATE: 2000-09-05
;   PRIOR APPLICATION NUMBER: 09/661,753
;   PRIOR FILING DATE: 2000-09-14
;   PRIOR APPLICATION NUMBER: 60/154,546
;   PRIOR FILING DATE: 1999-09-17
;   NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 28
;   LENGTH: 2745
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (842)...(2017)
US-09-948-002-28

```

Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	488	-GAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATT	546
Db	1317	GGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATT	1376
Qy	547	CCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCCGGAGTGGCTGTCTCT	606
Db	1377	CCTGGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTT	1436
Qy	607	TTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTC	666
Db	1437	TTGATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGAGGGGAAATTGAGGGCTTTC	1496
Qy	667	GCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACG	726
Db	1497	GCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACG	1556
Qy	727	GGTTCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTTC	786
Db	1557	GGTTCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTTC	1616
Qy	787	TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCC	846
Db	1617	TGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGAAAGCTCCCGGCACCGCC	1676
Qy	847	GAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACT	906
Db	1677	GA-----GCCCTGGACACCAACTATT	1697
Qy	907	GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGG	966
Db	1698	GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGG	1757
Qy	967	ACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGC	1026
Db	1758	ACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGC	1817
Qy	1027	CCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACC	1086
Db	1818	CCTGCCCCCTACATTTGGAGCCTGGACACGAGTACAGCAAGGTCTGGCCCTGTACAACC	1877
Qy	1087	AGCACAACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGC	1146
Db	1878	AGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGC	1937
Qy	1147	CCATCGTGTAATACTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1206
Db	1938	CCATCGTGTAATACTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1997
Qy	1207	GTTCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCC	1266
Db	1998	GCTCCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCAGGCCCCGGCCC	2057
Qy	1267	GGCCCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCC	1318
Db	2058	CACCCCGCCCCCGCCCGCCTTGCCATGGGGGCTGTATTTAAGGACACCGTGCCCC	2117

Qy 1319 AAGCCCACTTGGGATCGATTAAA 1341
 ||||| ||| | ||
 Db 2118 AAGCCCACTTGGGCCCCCATTAA 2140

RESULT 5

US-09-948-002-1

; Sequence 1, Application US/09948002

; Publication No. US20030050265A1

; GENERAL INFORMATION:

APPLICANT: Nicholas M. Dean

; APPLICANT: Susan F. Murray

TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH

; TITLE OF INVENTION: FACTOR BETA EXPRESSION

FILE REFERENCE: ISPH-0607

; CURRENT APPLICATION NUMBER: US/09/948,002

;
; CURRENT FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: 09/661,753

PRIOR FILING DATE: 2000-09-14

PRIOR APPLICATION NUMBER: 60/154,546

PRIOR FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 71

; SEQ ID NO 1

LENGTH: 2094

; TYPE: DNA

; ORGANISM: Mus musculus

FEATURE:

```

; NAME/KEY: CDS

```

LOCATION: (868) ... (2040)

US-09-948-002-1

Query Match 63.7%; Score 862.2; DB 11; Length 2094;
Best Local Similarity 81.7%; Pred. No. 6.6e-222;
Matches 1040; Conservative 0; Mismatches 188; Indels 45; Gaps 2;

Qy	10	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC	69
Db	863	CCCCCATGCCGCCCTCGGGGCTGCGGCTACTGCCGCTTCTGCTCCCACTCCCGTGGCTTC	922
Qy	70	TAGTGCTGACGCCTGGCCGGCCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGG	129
Db	923	TAGTGCTGACGCCCGGGAGGCCAGCCGCGGGACTCTCCACCTGCAAGACCATCGACATGG	982
Qy	130	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC	189
Db	983	AGCTGGTGAAACGGAAGCGCATCGAAGCCATCCGTGGCCAGATCCTGTCCAAACTAAGGC	1042
Qy	190	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	249
Db	1043	TCGCCAGTCCCCCAAGCCAGGGGGAGGTACCGCCCGGCCCGCTGCCCGAGGCGGTGCTCG	1102
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	309
Db	1103	CTTTGTACAACAGCACCCGCGACCGGGTGGCAGGCAGAGCGCCGACCCAGAGCCCGGAGC	1162
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCACCGCGTGCTAATGGTGGAAGCGGCAACC	369

Db	1163	CCGAAGCGGACTACTATGCTAAAGAGGTCACCCGCGTGCTAATGGTGGACCGCAACAACG	1222
Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	1223	CCATCTATGAGAAAACCAAAGACATCTCACACAGTATATATATGTTCTTCAATACGTCAG	1282
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	1283	ACATTCCGGGAAGCAGTGCCCGAACCCCATTTGCTGTCCCGTGCAGAGCTGCGCTTGCA	1342
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAAATACAGCAATGATTCCT	549
Db	1343	GATTAATAATCAAGTGTGGAGCAACATGTGGAACCTCTACCAGAAAATATAGCAACAATTCCT	1402
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	609
Db	1403	GGCGTTACCTTGGTAACCGGCTGCTGACCCCACTGATACGCCTGAGTGGCTGCTTTTG	1462
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	669
Db	1463	ACGTCACCTGGAGTTGTACGGCAGTGGCTGAACCAAGGAGACGGAATACAGGGCTTTCGAT	1522
Qy	670	TCAGTGCCCACTGTTCCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1523	TCAGCGCTCACTGCTCTTGTGACAGCAAAGATAACAACTCCACGTGGAAATCAACGGGA	1582
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	789
Db	1583	TCAGCCCCAAACGTCGGGGCGACCTGGGCACCATCCATGACATGAACCGGCCCTTCCTGC	1642
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1643	TCCTCATGGCCACCCCCCTGGAAAGGGCCCAGCACCTGCACAGCTCACGGCACCGGAGA-	1701
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1702	-----GCCCTGGATACCAACTATTGCT	1723
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCCGGAAGGACC	969
Db	1724	TCAGCTCCACAGAGAAGAACTGCTGTGTGCGGCAGCTGTACATTGACTTTAGGAAGGACC	1783
Qy	970	TGGGCTGGAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1784	TGGGTTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGTCTGGGACCT	1843
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1844	GCCCCATATATTGGAGCCTGGACACACAGTACAGCAAGGTCCTTGCCCTCTACAACCAAC	1903
Qy	1090	ACAACCCGGGCGCGTCCGGCGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149
Db	1904	ACAACCCGGGCGCTTCGGCGTCAACGTGCTGCGTGCCGAGGCTTGGAGCCACTGCCCA	1963
Qy	1150	TCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1209

```

Db          1964 TCGTCTACTACGTGGGTCGCAAGCCCAAGGTGGAGCAGTTGTCCAACATGATTGTGCGCT 2023
Qy          1210 CCTGCAAGTG CAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC 1269
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db          2024 CCTGCAAGTG CAGCTGAAGCCCCGCCCGC-----CCC GCCCTCCCGGCAGGCCCGGC 2077
Qy          1270 CCCACCCCGCCC 1282
           ||| | | | | | | |
Db          2078 CCCGCCCGCCC 2090

```

RESULT 6

US-09-948-002-27

; Sequence 27, Application US/09948002

; Publication No. US20030050265A1

; GENERAL INFORMATION:

; APPLICANT: Nicholas M. Dean

APPLICANT: Susan F. Murray

TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH

; TITLE OF INVENTION: FACTOR BETA EXPRESSION

FILE REFERENCE: ISPH-0607

; CURRENT APPLICATION NUMBER: US/09/948,002

; CURRENT FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: 09/661,753

; PRIOR FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 60/154,546

; PRIOR FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 71

; SEQ ID NO 27

; LENGTH: 1585

; TYPE: DNA

; ORGANISM: Rattus norvegicus

FEATURE:

; NAME/KEY: CDS

; LOCATION: (413) ... (1585)

US-09-948-002-27

Query Match 62.3%; Score 842.6; DB 11; Length 1585;

Best Local Similarity 82.1%; Pred. No. 1.2e-216;

Matches 999; Conservative 0; Mismatches 179; Indels 39; Gaps 1;

Qy	10	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC	69
Db	408	CCCCCATGCGGCCCTCGGGGCTGCGGCTGCTGCCGCTTCTGCTCCCACTCCCGTGGCTTC	467
Qy	70	TAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG	129
Db	468	TAGTGCTGACGCCCGGAGGCCAGCCGCGGGA CTCTCCACCTGCAAGACCATCGACATGG	527
Qy	130	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGGGCCAGATTCTGTCCAAGCTTCGGC	189
Db	528	AGCTGGTGA AACCGGAAGCGCATCGAAGCCATCCGTGGCCAGATCCTGTCCAAACTAAGGC	587
Qy	190	TTGCCAGCCCCCGAGCCAGGGGGAGCTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	249
Db	588	TCGCCAGTCCCCCGAGCCAGGGGGAGGTACCGCCGGGGCCCGCTGCCCGAGGCGGTGCTCG	647

Qy	250	CTCTTTTACAACAGTATCCCGCGACCGGGTAGCCGGGGAAAGTGTGCGAACCCGGAGCCCGAGC	309
Db	648	CTTTGTACAACAGCACCCGCGACCGGGTGGCAGGCGAGAGCGCTGACCCGGAGCCCGAGC	707
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	369
Db	708	CCGAGGCGGACTACTACGCCAAAGAAGTCACCCGCGTGCTAATGGTGGAACCGCAACAACG	767
Qy	370	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	429
Db	768	CAATCTATGACAAAACCAAAGACATCACACACAGTATATATATGTTCTTCAATACGTCAG	827
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	828	ACATTCCGGGAAGCAGTGCCAGAACCCCCATTGCTGTCCCGTGACAGAGCTGCGCCTGCAGA	887
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	549
Db	888	GATTCAAGTCAACTGTGGAGCAACACGTAGAACTCTACCAGAAATATAGCAACAATTCT	947
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	609
Db	948	GGCGTTACCTTGGTAACCGGCTGCTGACCCCCACTGATACGCCTGAGTGGCTGTCTTTG	1007
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGACAGAGGGCTATAGAGGGTTTTCGCC	669
Db	1008	ACGTCACTGGAGTTGTCCGGCAGTGGCTGAACCAAGGAGACGGAATACAGGGCTTTTCGCT	1067
Qy	670	TCAGTGCCCACTGTTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	1068	TCAGTGCTCACTGCTCTTGTGACAGCAAAGATAATGTACTCCACGTGGAAATCAATGGGA	1127
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCTCTGC	789
Db	1128	TCAGTCCCAAACGTCGAGGTGACCTGGGCACCATCCATGACATGAACCGACCCTTCTCTGC	1187
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	1188	TCCTCATGGCCACCCCCCTGGAAAGGGCTCAACACCTGCACAGCTCCAGGCACCGGAGA-	1246
Qy	850	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	909
Db	1247	-----GCCCTGGATACCAACTACTGCT	1268
Qy	910	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	969
Db	1269	TCAGCTCCACAGAGAAGAACTGCTGTGTACGGCAGCTGTACATTGACTTTAGGAAGGACC	1328
Qy	970	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1029
Db	1329	TGGGTTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGTCTGGGGCCCT	1388
Qy	1030	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1089
Db	1389	GCCCCATACATTTGGAGCCTGGACACACAGTACAGCAAGGTCCTTGCCCTCTACAACCAAC	1448
Qy	1090	ACAACCCGGGCGCGTGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1149

```

Db      1449  ACAAACCCGGGTGCTTCCGCATCACCGTGCTGCGTGCCGCAGGCTTTGGAGCCACTGCCCA 1508
Qy      1150  TCGTGTA CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1209
Db      1509  TCGTCTACTACGTGGGTGCGAAGCCCAAGGTGGAGCAGTTGTCCAACATGATCGTGCGCT 1568
Qy      1210  CCTGCAAGTGCAGCTGA 1226
Db      1569  CCTGCAAGTGCAGCTGA 1585

```

RESULT 7

US-09-756-283A-19

```
; Sequence 19, Application US/09756283A
; Patent No. US20020151478A1
; GENERAL INFORMATION:
;   APPLICANT: Chernajovsky, Yuti
;   APPLICANT: Dreja, Hanna Stina
;   APPLICANT: Adams, Gillian
;   TITLE OF INVENTION: Latent Fusion Protein
;   FILE REFERENCE: 0623.1000000
;   CURRENT APPLICATION NUMBER: US/09/756,283A
;   CURRENT FILING DATE: 2001-01-09
;   NUMBER OF SEQ ID NOS: 100
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
;   LENGTH: 1376
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: LAP-mIFNbeta construct
;   NAME/KEY: CDS
;   LOCATION: (1)..(1368)
US-09-756-283A-19
```

Query Match 48.6%; Score 657.2; DB 10; Length 1376;
Best Local Similarity 88.3%; Pred. No. 8.3e-167;
Matches 726; Conservative 0; Mismatches 93; Indels 3; Gaps 1;

Qy	15	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGTG	74
Db	1	ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTGTG	60
Qy	75	CTGACGCCTGGCCGGCCGGCCCGGAACTGTCCACCTGCAAGACCATCGACATGGAGCTG	134
Db	61	CTGACGCCTGGCCCCGCCGGCCCGGGAATATCCACCTGCAAGACTATCGACATGGAGCTG	120
Qy	135	GTGAAGCGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	194
Db	121	GTGAAGCGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	180
Qy	195	AGCCCCCGAGCCAGGGGGACGTGCCGCCC GGCCC GCTGCCTGAGGCAGTA CTGGCTCTT	254
Db	181	AGCCCCCGAGCCAGGGGGAGGTGCCGCCC GGCCC GCTGCCCAGGCCGTGCTCGCCCTG	240
Qy	255	TACAACAGTACCCGCGACCCGGGTAGCCGGGGAAAGTGTCAAACCGGAGCCCGAGCCAGAG	314


```

Db      1242 TACCGGCCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCTTTCCTGCTTCT 1301
Qy      794 CATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCT 834
      |||||
Db      1302 CATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCT 1342

```

RESULT 9

US-09-911-904-167

```

; Sequence 167, Application US/09911904
; Publication No. US20030096234A1
; GENERAL INFORMATION:
; APPLICANT: Farr, Spencer B.
; APPLICANT: Pickett, Gavin G.
; APPLICANT: Neft, Robin Eileen
; APPLICANT: Dunn, II, Robert Thomas
; TITLE OF INVENTION: CANINE TOXICITY GENES
; FILE REFERENCE: 400742000200
; CURRENT APPLICATION NUMBER: US/09/911,904
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/220,057
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(489)
; OTHER INFORMATION: n = A, T, C or G
US-09-911-904-167

```

```

Query Match          27.1%; Score 366.4; DB 11; Length 489;
Best Local Similarity 86.1%; Pred. No. 1.1e-88;
Matches 445; Conservative 0; Mismatches 32; Indels 40; Gaps 2;

```

```

Qy      778 GGCCCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCC 837
      | |||||
Db      1 GACCCCTTCCTGCTCCTCATGGCCACCCCACTGGAGAGGGCCCAGCACCTGCACAGCTCCC 60

Qy      838 GGCACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATA 897
      ||||| |||||
Db      61 GGCAGCGCCG-----GGCCCTGGACA 81

Qy      898 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 957
      |||||
Db      82 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 141

Qy      958 TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCT 1017
      |||||
Db      142 TCCGCAAGGATCTGGGCTGGAAGTGGATCCATGAGCCCAAGGGTTACCACGCTAACTTCT 201

Qy      1018 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1077
      |||||
Db      202 GCCTGGGGCCCTGCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC 261

```


Qy	1078	TGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	1137
Db	262	TGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGG	321
Qy	1138	AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	1197
Db	322	AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA	381
Qy	1198	TGATCGTGCCTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCG-CCCACAGCCCCGCCACC	1256
Db	382	TGATCGTGCCTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGTCGGCAGGCCCCGCCACC	441
Qy	1257	CGGCAGGCCCCGGCCCCACCCCGCCCGCTCACC	1293
Db	442	CGGCAGGNCCGGCCCCCGCCCGCTGCGCCGGG	478

RESULT 10

US-09-813-271B-1

; Sequence 1, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of
; biologically active protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. US20020115834A1artis Patent Department

; STREET: 564 Morris Avenue

; CITY: Summit

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07901

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,271B

; FILING DATE: 20-Mar-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP95/02719

; FILING DATE: 12-Jul-95

; APPLICATION NUMBER: EPO 94810439.3

; FILING DATE: 25-Jul-94

; ATTORNEY/AGENT INFORMATION:

; NAME: Pfeiffer, Hesna J. .

; REGISTRATION NUMBER: 22640

; REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908) 522-6940

; TELEFAX: (908) 522-6955

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

```

;           LENGTH: 339 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: double
;           TOPOLOGY: linear
;           MOLECULE TYPE: cDNA to mRNA
;           HYPOTHETICAL: NO
;           IMMEDIATE SOURCE:
;             CLONE: E. coli LC137/pPLMu.htGF-beta1 (DSM 5656)
;           FEATURE:
;             NAME/KEY: CDS
;             LOCATION: 1..336
;             OTHER INFORMATION: /product= "human TGF-beta1"
;           SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-813-271B-1

```

```

Query Match          22.8%;  Score 308.6;  DB 10;  Length 339;
Best Local Similarity 94.4%;  Pred. No. 3.5e-73;
Matches 320;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

```

```

Qy      888  GCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTC 947
          |||
Db      1    GCCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTG 60

Qy      948  TACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCAT 1007
          |||
Db      61   TACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCAT 120

Qy     1008  GCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAG 1067
          |||
Db     121   GCCAATTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAG 180

Qy     1068  GTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGCTCGGCGGCGCCGTGCTGCGTGCCG 1127
          |||
Db     181   GTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCG 240

Qy     1128  CAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAG 1187
          |||
Db     241   CAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAG 300

Qy     1188  CTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1226
          |||
Db     301   CTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 339

```

RESULT 11

US-09-906-158-3

; Sequence 3, Application US/09906158

; Publication No. US20030078217A1

; GENERAL INFORMATION:

; APPLICANT: Brett P. Monia

; APPLICANT: Susan M. Freier

; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR-BETA 3
EXPRESSION

; FILE REFERENCE: RTS-0257

; CURRENT APPLICATION NUMBER: US/09/906,158

; CURRENT FILING DATE: 2001-07-14

; NUMBER OF SEQ ID NOS: 168

; SEQ ID NO 3
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (254)...(1492)
US-09-906-158-3

Query Match 17.9%; Score 241.6; DB 11; Length 2574;
Best Local Similarity 53.8%; Pred. No. 5.9e-55;
Matches 670; Conservative 0; Mismatches 519; Indels 57; Gaps 6;

```
Qy      40  TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 99
      ||| ||| || ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      261 TGCACTTGCAAAGGGCTCTGGTGGTCTGGCCCTGCTGAACTTTGCCACGGTCAGCCTCT 320

Qy      100 GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 159
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      321 CTCTGTCCACTTGCAACACCTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGGAAGCCA 380

Qy      160 TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 219
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      381 TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTGATGA 440

Qy      220 CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 279
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      441 CCCACG-----TCCCTATCAGGTCTCTGGCCCTTTACAACAGCACCCGGGAGCTGCTGG 494

Qy      280 CCGGGGAAAAGTGTGCAACCGGAGCCCG-----AGCCAGAGGCGGACTACT 324
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      495 AGGAGATGCATGGGGAGAGGGAGGAAGGCTGCACCCAGGAAAAACCCGAGTCGGAATACT 554

Qy      325 ACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 384
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      555 ATGCCAAAGAAATCCATAAATTCGACATGATCCAGGGGCTGGCGGAGCACAAACGAACTGG 614

Qy      385 TCAAGGGCACCCCCCACAGCTTATATATGCTGTTCAACACGTCTGGAGCTCCGGGAAGCGG 444
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      615 CTGTCTGCCCTAAAGGAATTACCTCCAAGGTTTTCCGCTTCAATGTGTCTCTCAGTGGA 674

Qy      445 TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAG 504
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      675 AAAATAGAACCAACCTATTCCGAGCAGAATTCGGGTCTTGCGGGTGCCCAACCCAGCT 734

Qy      505 TGGAGCAGCACGTGGAG-----CTATACCAGAAATACAGCAATG 543
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      735 CTAAGCGGAATGAGCAGAGGATCGAGCTCTTCCAGATCCTTCGGCCAGATGAGCACATTG 794

Qy      544 ATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGCGTGT 603
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      795 CCAAACAGCGCTATATCGGTGGCAAGAATCTGCCACACGGGGCACTGCCGAGTGCGTGT 854

Qy      604 CCTTTGATGTCAACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTT 663
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db      855 CCTTTGATGTCACTGACACTGTGCGTGAGTGGCTGTTGAGAAGAGAGTCCAACCTAGGTC 914
```

```

Qy      664 TTCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACA-----CACTCCACGTGG 717
      |   |||   ||||| | ||| ||| |   ||   |||   |
Db      915 TAGAAATCAGCATTCACCTGTCCATGTACACCTTTTCAGCCCAATGGAGATATCCTGGAAA 974

Qy      718 AAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACC 777
      |   |||   ||| | | |   |   |   |||   ||   ||   |||   |
Db      975 ACATTCACGAGGTGATGGAAATCAAATTCAAAGGCGTGACAAATGAGGATGACCATGGCC 1034

Qy      778 GGCCCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCC 837
      |   |||   ||| | | |   |   |   |||   |||   |||   |||   |
Db     1035 GTGGAGATCTGGGGCGCCTCAAGAAGCAGAAGGATCACCACAACCCTCATCTAATCCTCA 1094

Qy      838 GG-----CACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAA---GG 888
      |   ||   |   |||   |||   |||   |||   |||   |||   |||   |||
Db     1095 TGATGATTCCCCCACACCGGCTCGACAACCCGGGCCAGGGGGGTGAGAGGAAGAAGCGGG 1154

Qy      889 CCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCT 948
      |   |||   ||| | ||||| ||| |   |||   |||||   |||||   |||||
Db     1155 CTTTGGACACCAATTACTGCTTCCGCAACTTGGAGGAGAACTGCTGTGTGCGCCCCCTCT 1214

Qy      949 ACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGAATTCATGAACCCAAGGGCTACCATG 1008
      |||||   |||||   |||||   |||||   |||||   |||||   |||||   |||||
Db     1215 ACATTGACTTCCGACAGGATCTGGGCTGGAAGTGGGTCCATGAACCTAAGGGCTACTATG 1274

Qy     1009 CCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGG 1068
      ||||   |||||   || || || || || || || || || || || || || || || ||
Db     1275 CCAACTTCTGCTCAGGCCCTTGCCCATACCTCCGAGTGACAGACACAACCCACAGCACGG 1334

Qy     1069 TCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCCGCGGCGCCGTGCTGCGTGCCGC 1128
      |   ||||   |||||   || || || || || || || || || || || || || || ||
Db     1335 TGCTGGGACTGTACAACACTCTGAACCCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCC 1394

Qy     1129 AGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCGCGCAAGCCCAAGGTGGAGCAGC 1188
      |||   |||||   ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1395 AGGACCTGGAGCCCCTGACCATCCTGTACTATGTTGGGAGGACCCCCAAAGTGGAGCAGC 1454

Qy     1189 TGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGC 1234
      |   |||||   || || || || || || || || || || || || || || || ||
Db     1455 TCTCCAACATGGTGGTGAAGTCTTGTAATGTAGCTGAGACCCAC 1500

```

RESULT 12

US-10-028-158-20

; Sequence 20, Application US/10028158

; Publication No. US20020110833A1

; GENERAL INFORMATION:

; APPLICANT: Caniggia, Isabella

; APPLICANT: Post, Martin

; APPLICANT: Lye, Stephen

; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF

; TITLE OF INVENTION: TROPHOBLAST

; FILE REFERENCE: 11757.38USWO

; CURRENT APPLICATION NUMBER: US/10/028,158

; CURRENT FILING DATE: 2001-12-20

; PRIOR APPLICATION NUMBER: US/09/380,662

; PRIOR FILING DATE: 1999-12-21

; PRIOR APPLICATION NUMBER: PCT/CA98/00180

```

; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (254)..(1492)
US-10-028-158-20

```

```

Query Match          17.9%;  Score 241.6;  DB 13;  Length 2574;
Best Local Similarity 53.8%;  Pred. No. 5.9e-55;
Matches 670;  Conservative 0;  Mismatches 519;  Indels 57;  Gaps 6;

```

```

Qy      40  TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCGCCG 99
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      261  TGCACTTGCAAAGGGCTCTGGTGGTCTGGCCCTGCTGAACTTTGCCACGGTCAGCCTCT 320

Qy      100  GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 159
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      321  CTCTGTCCACTTGCAACACCTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGAAGCCA 380

Qy      160  TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 219
      ||  ||  ||  |||||  ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      381  TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTGATGA 440

Qy      220  CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTTACAACAGTACCCGCGACCGGGTAG 279
      ||  ||  ||  |||||  ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      441  CCCACG-----TCCCTATCAGGTCTTGGCCCTTTTACAACAGCACCCGGGAGCTGCTGG 494

Qy      280  CCGGGGAAAGTGTGCAACCGGAGCCCG-----AGCCAGAGGCGGACTACT 324
      ||  ||  ||  ||  |||||  ||  |||||  |||||  |||||  |||||  |||||  |||||
Db      495  AGGAGATGCATGGGGAGAGGGAGGAAGGCTGCACCCAGGAAAAACCCGAGTCGGAATACT 554

Qy      325  ACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 384
      ||  |||||  ||  ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      555  ATGCCAAAGAAATCCATAAATTCGACATGATCCAGGGGCTGGCGGAGCACAACGAACTGG 614

Qy      385  TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG 444
      ||  ||  ||  ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      615  CTGTCTGCCCTAAAGGAATTACCTCCAAGGTTTCCGCTTCAATGTGTCTCAGTGGAGA 674

Qy      445  TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAG 504
      ||  ||  ||  ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      675  AAAATAGAACCAACCTATTCCGAGCAGAATTCCGGGTCTTGCGGGTGCCCAACCCAGCT 734

Qy      505  TGGAGCAGCACGTGGAG-----CTATACCAGAAATACAGCAATG 543
      ||||  ||  ||  |||||  |||||  |||||  |||||  |||||  |||||
Db      735  CTAAGCGGAATGAGCAGAGGATCGAGCTCTTCCAGATCCTTCGGCCAGATGAGCACATTG 794

Qy      544  ATTCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGT 603
      |||||  ||  ||  ||  |||||  |||||  |||||  |||||  |||||  |||||
Db      795  CCAAAACAGCGCTATATCGGTGGCAAGAATCTGCCACACGGGGCACTGCCGAGTGGCTGT 854

```



```

; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: EP 99116533
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 4382
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector for transforming supporting cell with a foreign to
express
; OTHER INFORMATION: a gene product of interest
US-09-957-458B-9

```

```

Query Match          17.5%; Score 236.4; DB 12; Length 4382;
Best Local Similarity 53.0%; Pred. No. 1.7e-53;
Matches 662; Conservative 0; Mismatches 531; Indels 57; Gaps 5;

```

```

Qy      30 CTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGG 89
      |||| || | |||   |||   | | ||| |   |||   |||||   ||
Db      455 CTGCAGCCCATGCACTTGCAAAGGGCTCTGGTAGTCCCTGGCCCTGCTGAACTTGCCACA 514

Qy      90 CCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGC 149
      | | | ||||| ||| || | ||| | | | | ||| ||| ||| |
Db      515 ATCAGCCTCTCTCTGTCCACTTGCAACCGTTGGACTTCGGCCACATCAAGAAGAAGAGG 574

Qy     150 ATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAG 209
      | || ||||| | || ||||| || ||||| ||| ||||| ||||| |
Db      575 GTGGAAGCCATTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCA 634

Qy     210 GGGGACGTGCCGCCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGC 269
      ||  || | | ||  || || | ||||| ||||| ||||| |||||
Db      635 TCGGTGATGACCCACG-----TCCCCATCAGGTCTGGCACTTTACAACAGCACCCGG 688

Qy     270 GACCGGGTAGCCG-----GGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG 314
      ||  | | | | |||| | | | | | | | | |||
Db      689 GAGTTGCTGGAAGAGATGCACGGGGAGAGGGAGGAAGGCTGCACTCAGGAGACCTCGGAG 748

Qy     315 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 374
      | || ||||| ||||| ||| || | ||| | | | | | | | |
Db      749 TCTGAGTACTATGCCAAAGAGATCCATAAATTCGACATGATCCAGGGACTGGCGGAGCAC 808

Qy     375 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC 434
      |||| | | | ||| || | ||| | | | | | | | |
Db      809 AATGAACTGGCCGTCTGCCCAAAGGAATTACCTCTAAGGTTTTTCGTTTCAATGTGTCC 868

Qy     435 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 494
      | | | | | | | | | | | | | | | | | | | | | |
Db      869 TCAGTGAGAAAAAATGGAACCAATCTGTTCCGGGCAGAGTTCCGGGTCTTGCGGGTGCCC 928

Qy     495 A-----AGTTAAAAGTGAGCAGCACGTGGAGCTATAACCAGAAA 533
      | | | | | | | | | | | | | | | | | |
Db      929 AACCCAGCTCCAAGCGCACAGAGCAGAGAATTGAGCTCTTCCAGATACTTCGACCGGAT 988

Qy     534 TACAGCAATGATTCTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCG 593
      | || | ||||| | | | | | | | | | | | |

```

```

RESULT 14
US-09-906-158-10
; Sequence 10, Application US/09906158
; Publication No. US20030078217A1
; GENERAL INFORMATION:
;   APPLICANT: Brett P. Monia
;   APPLICANT: Susan M. Freier
;   TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR-BETA 3
EXPRESSION
;   FILE REFERENCE: RTS-0257

```


; CURRENT APPLICATION NUMBER: US/09/906,158
; CURRENT FILING DATE: 2001-07-14
; NUMBER OF SEQ ID NOS: 168
; SEQ ID NO 10
; LENGTH: 2879
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (611) ... (1843)
US-09-906-158-10

Query Match 17.2%; Score 233.2; DB 11; Length 2879;
Best Local Similarity 52.9%; Pred. No. 1.1e-52;
Matches 657; Conservative 0; Mismatches 528; Indels 57; Gaps 5;

```
Qy      40 TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 99
      |||  |||  | | ||| | |||  |||  |||  |||  |||  |||  |||
Db      612 TGCACTTGCAAAGGGCTCTGGTAGTCTTGGCCCTGCTGAACTTGGCCACAATCAGCCTCT 671

Qy      100 GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 159
      ||||| |||  ||  |||  |||  |||  |||  |||  |||  |||  |||
Db      672 CTCTGTCCACTTGCAACACGTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGGAAGCCA 731

Qy      160 TTCGCGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 219
      ||  ||  |||  ||  |||  |||  |||  |||  |||  |||  ||  ||
Db      732 TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCATCGGTGATGA 791

Qy      220 CGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 279
      | | ||  | ||  | |||  |||  |||  |||  |||  |||  |||  |||
Db      792 CCCACG-----TCCCCATCAGGTCTTGGCACTTTTACAACAGCACCCGGGAGTTGCTGG 845

Qy      280 CCG-----GGGAAAGTGTGCAACCGGAGCCCGAGCCAGAGGCGGACTACT 324
      |  |||  ||  ||  ||  |||  |||  |||  |||  |||  |||
Db      846 AAGAGATGCACGGGGAGAGGGAGGAAGGCTGCACTCAGGAGACCTCGGAGTCTGAGTACT 905

Qy      325 ACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 384
      | |||| |||  ||  |||  |||  ||  ||  ||  |||  |||
Db      906 ATGCCAAAGAGATCCATAAATTCGACATGATCCAGGGACTGGCGGAGCACAATGAACTGG 965

Qy      385 TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG 444
      |  ||  |||  |  |||  |||  |||  ||  ||  ||  |||  |||
Db      966 CCGTCTGCCCCAAAGGAATTACCTCTAAGGTTTTTCGTTTCAATGTGTCTCAGTGGAGA 1025

Qy      445 TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCA----- 495
      | | |  |||  ||  ||  ||  ||  ||  |||  |||  |||  |||
Db      1026 AAAATGGAACCAATCTGTTCCGGGCAGAGTTCCGGGTCTTGCGGGTGCCCAACCCAGCT 1085

Qy      496 -----AGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATG 543
      ||  ||  |||  |||  ||  ||  ||  |||  |||  |||  |||
Db      1086 CCAAGCGCACAGAGCAGAGAATTGAGCTCTTCCAGATACTTCGACCGGATGAGCACATAG 1145

Qy      544 ATTCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGT 603
      ||||| ||  ||  ||  |||  |||  |||  |||  |||  |||  |||
Db      1146 CCAAGCAGCGCTACATAGGTGGCAAGAATCTGCCCAAGGGGCACCGCTGAATGGCTGT 1205

Qy      604 CCTTTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTT 663
```

Db	1206	CTTTCGATGTCACTGACACTGTGCGCGAGTGGCTGTTGAGGAGAGAGTCCAACCTGGGTC	1265
Qy	664	TTCGCCTCAGTGCCCACTGTTCCCTGTGACA-----GCAAAGATAACACACTCCACG	714
Db	1266	TGGAAATCAGCATCCACTGTCCATGTACACCTTTTCAGCCCAATGGAGACATACTGGAAA	1325
Qy	715	TGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGA	774
Db	1326	ATGTTTCATGAGGTGATGGAAATCAAATTCAAAGGAGTGGACAATGAAGATGACCATGGCC	1385
Qy	775	ACCGGCCCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCT	834
Db	1386	GTGGAGACCTGGGGCGTCTCAAGAAGCAAAAGGATCACCACAACCCACACCTGATCCTCA	1445
Qy	835	CCCGGCACCGCCGAGCCCTGGATACCAACGACTACAAGG-----ATGACGACGACAAGG	888
Db	1446	TGATGATCCCCCACACCGACTGGACAGCCAGGCCAGGGCAGTCAGAGGAAGAAGAGGG	1505
Qy	889	CCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCT	948
Db	1506	CCCTGGACACCAATTACTGCTTCCGCAACCTGGAGGAGAAGTCTGTGTACGCCCCCTTT	1565
Qy	949	ACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATG	1008
Db	1566	ATATTGACTTCCGGCAGGATCTAGGCTGGAAATGGGTCCACGAACCTAAGGGTTACTATG	1625
Qy	1009	CCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGG	1068
Db	1626	CCAACTTCTGCTCAGGCCCTTGCCCATACCTCCGCAGCGCAGACACAACCCATAGCACGG	1685
Qy	1069	TCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGC	1128
Db	1686	TGCTTGGACTATACAACACCCTGAACCCAGAGGCGTCTGCCTCGCCATGCTGCGTCCCCC	1745
Qy	1129	AGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGC	1188
Db	1746	AGGACCTGGAGCCCCTGACCATCTGTACTATGTGGGCAGAACCCCCAAGGTGGAGCAGC	1805
Qy	1189	TGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCC	1230
Db	1806	TGTCCAACATGGTGGTGAAGTCGTGTAAGTGCAGCTGAGGCC	1847

RESULT 15

US-09-813-271B-7

; Sequence 7, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of
; biologically active protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. US20020115834A1artis Patent Department

; STREET: 564 Morris Avenue

```

;          CITY: Summit
;          STATE: New Jersey
;          COUNTRY: USA
;          ZIP: 07901
;
;  COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;
;  CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/09/813,271B
;          FILING DATE: 20-Mar-2001
;
;  PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: PCT/EP95/02719
;          FILING DATE: 12-Jul-95
;          APPLICATION NUMBER: EPO 94810439.3
;          FILING DATE: 25-Jul-94
;
;  ATTORNEY/AGENT INFORMATION:
;          NAME: Pfeiffer, Hesna J. .
;          REGISTRATION NUMBER: 22640
;          REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN
;
;  TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (908) 522-6940
;          TELEFAX: (908) 522-6955
;
;  INFORMATION FOR SEQ ID NO: 7:
;          SEQUENCE CHARACTERISTICS:
;              LENGTH: 336 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: double
;              TOPOLOGY: linear
;
;  MOLECULE TYPE: other nucleic acid
;          DESCRIPTION: /desc = "recombinant hybrid DNA of
;
;  IMMEDIATE SOURCE:
;          CLONE: E. coli LC137/pPLMu.TGF-beta1(44/45)beta3
;
;  FEATURE:
;          NAME/KEY: mat_peptide
;          LOCATION: 1..132
;          OTHER INFORMATION: /product= "N-terminal 44 amino
;              acids of human TGF-beta1"
;
;  FEATURE:
;          NAME/KEY: mat_peptide
;          LOCATION: 133..336
;          OTHER INFORMATION: /product= "C-terminal 68 amino
;              acids of human TGF-beta3"
;
;  FEATURE:
;          NAME/KEY: CDS
;          LOCATION: 1..336
;          OTHER INFORMATION: /product= "hybrid TGF-beta named
;              TGF-beta1-3"
;
;  SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-271B-7

```

```

Query Match          16.8%;  Score 227.2;  DB 10;  Length 336;
Best Local Similarity 79.8%;  Pred. No. 2.7e-51;
Matches 268;  Conservative 0;  Mismatches 68;  Indels 0;  Gaps 0;

```

Qy 888 GCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTC 947

2: em_esthum:*
 3: em_estin:*
 4: em_estmu:*
 5: em_estov:*
 6: em_estpl:*
 7: em_estro:*
 8: em_htc:*
 9: gb_est1:*
 10: gb_est2:*
 11: gb_htc:*
 12: gb_est3:*
 13: gb_est4:*
 14: gb_est5:*
 15: em_estfun:*
 16: em_estom:*
 17: em_gss_hum:*
 18: em_gss_inv:*
 19: em_gss_pln:*
 20: em_gss_vrt:*
 21: em_gss_fun:*
 22: em_gss_mam:*
 23: em_gss_mus:*
 24: em_gss_pro:*
 25: em_gss_rod:*
 26: em_gss_phg:*
 27: em_gss_vrl:*
 28: gb_gss1:*
 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result			%		Query				Description
	No.	Score	Match	Length	DB	ID			
	1	718.8	53.1	1072	12	BM562135		BM562135 AGENCOURT	
c	2	713.6	52.7	1201	13	BX355682		BX355682 BX355682	
c	3	702.6	51.9	983	13	BX335351		BX335351 BX335351	
	4	693.2	51.2	900	13	BX349319		BX349319 BX349319	
c	5	661.2	48.9	1041	13	BX324511		BX324511 BX324511	
c	6	622.4	46.0	888	13	BX434425		BX434425 BX434425	
	7	622.2	46.0	1043	12	BM555996		BM555996 AGENCOURT	
c	8	615.2	45.5	950	13	BX383773		BX383773 BX383773	
	9	580.6	42.9	859	9	AL530081		AL530081 AL530081	
	10	573.8	42.4	713	10	BE312000		BE312000 601154768	
	11	565.8	41.8	717	10	BE260971		BE260971 601153715	
	12	564.8	41.7	902	13	BQ675698		BQ675698 AGENCOURT	
	13	537.6	39.7	925	12	BI818841		BI818841 603037307	
c	14	537	39.7	841	9	AL530080		AL530080 AL530080	
c	15	536.6	39.7	956	12	BI084718		BI084718 602869722	
	16	521.8	38.6	1093	12	BQ054305		BQ054305 AGENCOURT	
	17	517	38.2	871	13	BQ952138		BQ952138 AGENCOURT	
	18	514	38.0	1049	12	BQ053417		BQ053417 AGENCOURT	
c	19	508.4	37.6	757	14	CA309731		CA309731 UI-H-FT1-	

20	506.6	37.4	1013	12	BM469326	BM469326	AGENCOURT
21	504.8	37.3	943	12	BI909079	BI909079	603070060
22	503.4	37.2	773	12	BI195242	BI195242	602944518
23	502	37.1	821	10	BG748049	BG748049	602705544
c 24	501.2	37.0	1047	13	BX325288	BX325288	BX325288
c 25	481.4	35.6	727	9	AI131171	AI131171	qc15c11.x
c 26	476.2	35.2	785	9	AI148173	AI148173	qb56d01.x
27	476	35.2	675	10	BE261764	BE261764	601147839
28	475	35.1	637	12	BG938856	BG938856	cn28g07.y
29	471	34.8	748	10	BF682938	BF682938	602117568
c 30	465.8	34.4	722	13	BU632686	BU632686	UI-H-FE1-
31	464.2	34.3	960	13	BU957001	BU957001	AGENCOURT
c 32	464	34.3	928	13	BX412434	BX412434	BX412434
c 33	463.8	34.3	695	14	CA425775	CA425775	UI-H-FE1-
34	462.8	34.2	659	10	BG500488	BG500488	602544826
35	462	34.1	565	14	CD287835	CD287835	1_M15.abd
c 36	461.8	34.1	697	13	BU633535	BU633535	UI-H-FL1-
c 37	461.2	34.1	700	13	BU633922	BU633922	UI-H-FL1-
c 38	460.6	34.0	690	13	BU625377	BU625377	UI-H-FG1-
c 39	460.6	34.0	697	14	CA426391	CA426391	UI-H-FE1-
c 40	459.8	34.0	778	9	AI743724	AI743724	wg53b06.x
41	459.4	34.0	814	12	BI686959	BI686959	603313837
c 42	457.2	33.8	1015	10	BG683840	BG683840	602651694
43	456.8	33.8	599	12	BI905963	BI905963	603062849
c 44	454.4	33.6	720	14	CA416525	CA416525	UI-H-FE0-
45	452.2	33.4	548	12	BM068788	BM068788	ie85a02.y

ALIGNMENTS

RESULT 1
BM562135

LOCUS BM562135 1072 bp mRNA linear EST 20-FEB-2002

DEFINITION AGENCOURT_6562032 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745463
5', mRNA sequence.

ACCESSION BM562135

VERSION BM562135.1 GI:18807966

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1072)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12768 row: h column: 08

High quality sequence start: 9
High quality sequence stop: 647.

FEATURES Location/Qualifiers
 source 1. .1072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5745463"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_118"
 /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."
BASE COUNT 217 a 356 c 313 g 186 t
ORIGIN

Query Match 53.1%; Score 718.8; DB 12; Length 1072;
Best Local Similarity 82.7%; Pred. No. 1.4e-146;
Matches 888; Conservative 0; Mismatches 142; Indels 44; Gaps 4;

Qy 131 GCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCT 190
 |||||
Db 20 GCTGGTGAAGCGGAAGCGCATCTAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCT 79

Qy 191 TGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGC 250
 |||||
Db 80 CGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGC 139

Qy 251 TCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCCGAGCC 310
 || |||||
Db 140 CCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCCGAGCC 199

Qy 311 AGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCA 370
 |||||
Db 200 TGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGA 259

Qy 371 AATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGA 430
 |||||
Db 260 AATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGA 319

Qy 431 GCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAG 490
 |||||
Db 320 GCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAG 379

Qy 491 GCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTG 550
 |||||
Db 380 GCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTG 439

Qy 551 GCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTGA 610
 |||

```

Db      440 GCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGA 499
Qy      611 TGTACACCGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCT 670
        |||
Db      500 TGTACACCGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCT 559
Qy      671 CAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTT 730
        |||
Db      560 TAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTT 619
Qy      731 CAATTCTGGCCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCT 790
        |||
Db      620 CACTACCGGCCGCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCT 679
Qy      791 CCTCATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGC 850
        |||
Db      680 TCTCATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGC 739
Qy      851 CCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTT 910
        |||
Db      740 -----CCCTGGACACCAACTATTGCTT 761
Qy      911 CAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCT 970
        |||
Db      762 CAGCTCCCGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCGAGGACCT 821
Qy      971 GGGCTGGAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTG 1030
        |||
Db      822 CGGCTGGAAGTGGATCCACGAGCCCCAAGGCTACCATGCCCACTTCTGCCTCGGGCCCTG 881
Qy      1031 TCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACAGCA 1090
        |||
Db      882 CCCCTACATTTGGAGCCTGGACACGCGTACCAGCAAGTCCCTGGCCCTGTACACCCGCCA 941
Qy      1091 CAACCCGGGCGCGTCGGCGGCGCCGTGCTGC--GTGCCGCGAGGCGCTGGAGCCACTGCCC 1148
        |||
Db      942 TAACCGGGCGCCCTCGGCGGCGGGTTTGCTGGGGCCCCAGGCCCTGGACCCCTGCCC 1001
Qy      1149 AT--CGTGTA TACTACGTGGGCGCGC--AAGCCCCAAGGTGGAGCAGCTGTCCAACAT 1198
        |||
Db      1002 CTTGCGGGTACTACCTGGGGCGCCAAGCCCCAAGTGGGAACACCTGTCCACAT 1055

```

RESULT 2

BX355682/c

LOCUS BX355682 1201 bp mRNA linear EST 05-MAY-2003

DEFINITION BX355682 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI002YJ05 3-PRIME, mRNA sequence.

ACCESSION BX355682

VERSION BX355682.1 GI:30371987

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1201)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

clone CS0DI013YH16 3-PRIME, mRNA sequence.

ACCESSION BX335351

VERSION BX335351.1 GI:30308367

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 983)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI013DD08NP1&cluster=9160.r>. Contact : Feng Liang Email : fliang@lifetech.com

URL : <http://fulllength.invitrogen.com/> InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DI013DD08NP1.

FEATURES Location/Qualifiers

source 1..983

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DI013YH16"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 170 a 280 c 323 g 203 t 7 others

ORIGIN

Query Match 51.9%; Score 702.6; DB 13; Length 983;

Best Local Similarity 84.5%; Pred. No. 4.7e-143;

Matches 859; Conservative 5; Mismatches 105; Indels 47; Gaps 5;

Qy	282	GGGGAAAGTGTCTGAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACC	341
Db	983	GGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCGGACTACTACGCCAAGGAGGTCACC	924
Qy	342	CGCGTGCTAATGGTGGAAAAGCGGCAACCAATCTATGATAAAATTCAAGGGCACCCCCAC	401
Db	923	CGCGTGCTAATGGTGGARACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACACAC	864
Qy	402	AGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTG	461
Db	863	AGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGTTG	804
Qy	462	CTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAG	521
Db	803	CTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAG	744

Qy	522	CTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCC	581
Db	743	CTGTACCAGAAATACAGCAACAATYCCTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	684
Qy	582	AGTGA CTCA CCGGAGTGGCTGTCTTTGATGTCA CCGGAGTTGTGCGGCAGTGGCTGACC	641
Db	683	AGCGACTCGCCAGAGTGGTTATCTTTTGATGTCA CCGGAGTTGTGCGGCAGTGGTTGAGC	624
Qy	642	CGCAGAGAGGCTATAGAGGGTTTTCGCCTCAGTGCCCACTGTTCTGTGACAGCAAAGAT	701
Db	623	CGTGGAGGGGAAATTGA-GGCTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGAT	565
Qy	702	AACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACC	761
Db	564	AACACACTGCAAGTGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACC	505
Qy	762	ATTCACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAG	821
Db	504	ATTCATGGCATGAACCGGCCCTTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAG	445
Qy	822	CACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGAC	881
Db	444	CATCTGCAAAGCTCCCGGCACCGCCGA-----	418
Qy	882	GACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	941
Db	417	-----GCCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	364
Qy	942	CAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGC	1001
Db	363	CAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGC	304
Qy	1002	TACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTAC	1061
Db	303	TACCATGCCAACTTCTGCCTCGGGCCCTGCCCTACATTTGGAGCCTGGACACGCAGTAM	244
Qy	1062	AGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGC	1121
		: :	
Db	243	MMCAAGGTCCTGGCCMTGTACAACCAGCATAACCC-GGCGCCTCGGCGGCGCCGTNGT-C	186
Qy	1122	GTGCCG CAGGCGCTGGAGCCACTGCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTG	1181
Db	185	GTGCCG CAGGCGCTGGAGCCGCTGCCCATCGTGTA CTACGTGGGCTGCAAGCCCAAGGTG	126
Qy	1182	GAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCC	1241
Db	125	GAGCAGCTGTCCAACATGATCGTGCGTCTCTGCAAGTGCAGCTGAGGTCCCGCCCCGCC-	67
Qy	1242	ACAGCCCCGCCCAACCGGCAGGCCCGGCCCAACCCCGCCCGCCTCACCGGGGCTG	1297
Db	66	----CCGCCCCGCCCGGCAGGCCCGGCCCAACCCCGCCCGCCTGCCTTG	15

RESULT 4
BX349319
LOCUS

BX349319

900 bp

mRNA

linear

EST 05-MAY-2003

DEFINITION BX349319 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0DL010YL07 5-PRIME, mRNA sequence.
 ACCESSION BX349319
 VERSION BX349319.1 GI:30379410
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 900)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9160.r For
 more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAG059ZD04_CS05596_1&cluster=9160.r)
[cgi-bin/cluster.cgi?seq=CS0BAG059ZD04_CS05596_1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAG059ZD04_CS05596_1&cluster=9160.r).
 Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAG059ZD04_CS05596_1.
 FEATURES Location/Qualifiers
 source 1. .900
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL010YL07"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 179 a 292 c 270 g 155 t 4 others
 ORIGIN

Query Match 51.2%; Score 693.2; DB 13; Length 900;
 Best Local Similarity 87.8%; Pred. No. 5.1e-141;
 Matches 755; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

Qy 10 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
 || ||| |||| || ||||||||||| ||||||||||| ||||||||||| ||
 Db 30 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 89
 Qy 70 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG 129
 | ||||||||||||||||||||||| ||||| ||||||||||| |||||||||||
 Db 90 TGGTGCTGACGCCTGGCCGGCCGGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG 149
 Qy 130 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
 ||||||||||||||||||||||| ||||||||||| ||||||||||| ||||
 Db 150 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 209

Qy	190	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG	249
Db	210	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG	269
Qy	250	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTGGAACCGGAGCCCGAGC	309
Db	270	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	329
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC	369
Db	330	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACG	389
Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG	429
Db	390	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	449
Qy	430	AGCTCCGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	450	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	509
Qy	490	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	549
Db	510	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT	569
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	609
Db	570	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	629
Qy	610	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	669
Db	630	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	689
Qy	670	TCAGTGCCCACTGTTCCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	729
Db	690	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	749
Qy	730	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCTCTGC	789
Db	750	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCCTTTCTCTGC	809
Qy	790	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	849
Db	810	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGNCGAG	869
Qy	850	CCCTGGATACCAACGACTAC	869
Db	870	CCCTGGNACACCACTATTGC	889

RESULT 5

BX324511/c

LOCUS BX324511 1041 bp mRNA linear EST 02-MAY-2003

DEFINITION BX324511 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DC024YD20 3-PRIME, mRNA sequence.

ACCESSION BX324511

VERSION BX324511.1 GI:30332381

Accession	Contig	Position	Sequence	Length
Db	786	727	CTGTACCAGAAATACAGCAACAATTCTCTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	727
Qy	582	641	AGTGACTCACCGGAGTGGCTGTCTTTGATGTACCCGGAGTTGTGCGGCAGTGGCTGACC	641
Db	726	667	AGCGACTCGCCAGAGTGGTTATCTTTTGATGTACCCGGAGTTGTGCGGCAGTGGTTGAGC	667
Qy	642	701	CGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTGTTCTCTGTGACAGCAAAGAT	701
Db	666	608	CGTGGAGGGGAAATTGA-GGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGAT	608
Qy	702	761	AACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACC	761
Db	607	548	AACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACC	548
Qy	762	821	ATTCACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAG	821
Db	547	488	ATTCATGGCATGAACCGGCCCTTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAG	488
Qy	822	881	CACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGAC	881
Db	487	452	CATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC-----	452
Qy	882	941	GACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	941
Db	451	407	-----ACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	407
Qy	942	1001	CAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGC	1001
Db	406	347	CAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGC	347
Qy	1002	1061	TACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTAC	1061
Db	346	287	TACCATGCCAATTTCTGCCTCGGGCCCTGCCCTACATTGGAGCCTGGACACGCAGTAC	287
Qy	1062	1121	AGCAAGGTCTTGCTCTGTACAACCAGCACAAACCCGGGCGCGTTCGGCGGCGCCGTGCTGC	1121
Db	286	229	AGCAAGGTCTTGCCCTGTACAACCAGCATAACCC-GGCGCCTCGGCGGCGCCGTG-TGC	229
Qy	1122	1181	GTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTG	1181
Db	228	169	GTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTG	169
Qy	1182	1241	GAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCC	1241
Db	168	109	GAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGTCCCCGCCCCGCC	109
Qy	1242	1293	ACAGCCCCGCCCCACCGGCAGGCCCGGCCACCCCCGCCCCGCT-----CACCGGG	1293
Db	108	49	CGCCCCGCCCCCGGCAGGCCCGGCCACCCCCGCCCCGCCCCGCTGCCTTGCCCATGGGG	49
Qy	1294	1341	GCTGTATTTAAGGACATCGTGCCCCAAGCCCACTGGGATCGATTAAA	1341
Db	48	1	GCTGTATTTAAGGACACCCGTCCCCAAGCCCACTGGSCACCCAYTAA	1

BX434425/c
 LOCUS BX434425 888 bp mRNA linear EST 15-MAY-2003
 DEFINITION BX434425 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE014YE16
 3-PRIME, mRNA sequence.
 ACCESSION BX434425
 VERSION BX434425.1 GI:30779291
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 888)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9160.r For
 more information about this cluster, see
[http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAK028AB08NM1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAK028AB08NM1&cluster=9160.r). Contact :
 Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAK028AB08NM1.
 FEATURES Location/Qualifiers
 source 1. 888
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE014YE16"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."
 BASE COUNT 156 a 253 c 302 g 176 t 1 others
 ORIGIN

Query Match 46.0%; Score 622.4; DB 13; Length 888;
 Best Local Similarity 83.6%; Pred. No. 1.5e-125;
 Matches 772; Conservative 0; Mismatches 102; Indels 49; Gaps 4;

Qy 432 CTCCGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGG 491
 ||| ||||| || ||||| | ||| | ||||| ||||| ||||| |||||
 Db 888 CTCCAAAAAGCGGTACCTGAACCCTGTTGTCTNCCCGGGCAGAGCTGCGTCTGCTGA-G 830
 Qy 492 CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCTGG 551
 ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 829 CTCAAGTTTAAATGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGG 770
 Qy 552 CGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGAT 611
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 769 CGATACCTCAGCAACCGGCTGCTGGCACACAGCGACTCGCCAGAGTGGTTATCTTTTGAT 710

Qy 612 GTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTC 671
|||||

Db 709 GTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTT 650

Qy 672 AGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTC 731
|||

Db 649 AGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTC 590

Qy 732 AATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTTGCTC 791
|||

Db 589 ACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTTCTTGCTT 530

Qy 792 CTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC 851
|||||

Db 529 CTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA--- 473

Qy 852 CTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTC 911
|||||

Db 472 -----GCCCTGGACACCAACTATTGCTTC 449

Qy 912 AGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTG 971
|||||

Db 448 AGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTC 389

Qy 972 GGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGT 1031
|||||

Db 388 GGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGC 329

Qy 1032 CCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCAC 1091
|||||

Db 328 CCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCAT 269

Qy 1092 AACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATC 1151
|||||

Db 268 AACCCGGGCGCCTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATC 209

Qy 1152 GTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCC 1211
|||||

Db 208 GTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTCC 149

Qy 1212 TGCAAGTGCAGCTGAGGCCCCGCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGCCC 1271
|||||

Db 148 TGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGGCAGGCCCCGGCCCCACCC 89

Qy 1272 CACCCCCGCCCCGCT-----CACCGGGCTGTATTTAAGGACA-TCGTGCCCCAAGC 1322
|||

Db 88 CGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCCGTGCCCCAAGC 29

Qy 1323 CCACTTGGGATCGATTAAAGCGG 1345
|||

Db 28 CCACCTGGGCCCCATTAAAGAGG 6

LOCUS BM555996 1043 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOURT_6544437 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5550039
 5', mRNA sequence.
 ACCESSION BM555996
 VERSION BM555996.1 GI:18796907
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1043)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12261 row: i column: 16
 High quality sequence stop: 702.
 FEATURES
 source Location/Qualifiers
 1..1043
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5550039"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_88"
 /note="Organ: small intestine; Vector: pCMV-SPORT6;
 Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
 oligo-dT primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 202 a 329 c 345 g 154 t 13 others
 ORIGIN

Query Match 46.0%; Score 622.2; DB 12; Length 1043;
 Best Local Similarity 82.2%; Pred. No. 1.7e-125;
 Matches 736; Conservative 0; Mismatches 120; Indels 39; Gaps 1;

Qy 400 ACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTAT 459
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1 ACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGT 60
 Qy 460 TGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG 519
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 61 TGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG 120
 Qy 520 AGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCC 579
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 121 AGCTGTACCAGAAATACAGCAACAATTCCTGGCGATACCTCAGCAACCGGCTGCTGGCAC 180

Qy	580	CCAGTGA	CTCACC	GGAGTG	GCTGTC	CTTTGAT	GTCA	CCGGAG	TTGTG	CGGCAG	TGGCTGA	639
Db	181	CCAGCGA	CTCGCC	AGAGTG	GTTATC	TTTGTAT	GTCA	CCGGAG	TTGTG	CGGCAG	TGGTTGA	240
Qy	640	CCCGCAG	AGAGG	CTATAG	AGGGT	TTTCGC	CCTCAG	TGCCCC	ACTGTT	CCTGTG	ACAGCAA	AG 699
Db	241	GCCGTGG	AGGGG	AAATTG	AGGGC	TTTCGC	CCTTAG	CGCCCC	ACTGCT	CCTGTG	ACAGCAG	GG 300
Qy	700	ATAACAC	ACTCC	ACGTGG	AAATTA	ACGGGT	TCAATT	CTGGCC	GCCGGG	TGACCT	TGGCCA	759
Db	301	ATAACAC	ACTGCA	AGTGG	ACATCA	ACGGGT	CACTAC	CGCCG	CCGAGG	TGACCT	TGGCCA	360
Qy	760	CCATTCA	CGGCAT	GAAAC	CGGCC	TTCTCT	GCTCCT	CATGGC	CACCCC	GCTGG	AGAGGG	CCC 819
Db	361	CCATTCA	TGGCAT	GAAAC	CGGCC	TTCTCT	GCTCCT	CATGGC	CACCCC	GCTGG	AGAGGG	CCC 420
Qy	820	AGCACCT	GCACAG	CTCCCG	GACCG	CCGAGC	CTTGGA	TACCA	ACGACT	ACAAGG	ATGACG	879
Db	421	AGCATCT	GCAAAG	CTCCCG	GACCG	CCGA-----						449
Qy	880	ACGACA	AGGCC	CTGGAT	ACCAACT	ACTGCT	TTCAGC	TCCACG	GAGAAG	AACTG	CTGCGT	GTC 939
Db	450	-----	GCCCTG	GACAC	CAACTA	TTCGCT	TTCAGC	TCCACG	GAGAAG	AACTG	CTGCGT	GTC 501
Qy	940	GGCAGCT	CTACAT	TGACTT	CCGGA	AGGAC	CTGGG	CTGGA	AGTGG	ATTCA	TGAAC	CCAAGG 999
Db	502	GGCAGCT	GTACAT	TGACTT	CCGCA	AGGAC	CTCGG	CTGGA	AGTGG	ATCCAC	GAGCC	CAAGG 561
Qy	1000	GCTACCA	TGCCAA	TTTCTG	CCTGGG	GGCCCT	GTCCCT	TACATC	TGGAGC	CTAGAC	ACTCAGT	1059
Db	562	GCTACCA	TGCCAA	CTTCTG	CCTCGG	GCCCTG	CCCC	TACATT	TGGAGC	CTGGAC	ACGCAGT	621
Qy	1060	ACAGCA	AGGTC	CCTGG	CTCTGT	ACAACC	AGCACA	ACCCGG	GCGCG	TGCGG	CGCGC	CGTGCT 1119
Db	622	ACAGCA	AGGTC	CCTGG	CCCTGT	ACAACC	AGCATA	ACCCGG	GCGCCT	CGGCG	GCGCG	CGTGCT 681
Qy	1120	GCGTG	CCG	CAGG	CGCTG	GAGC	CACTG	CCCAT	CGTGT	ACTAC	GTGGG	CCGCAAGCCCAAGG 1179
Db	682	GCGTG	CCG	CAGG	CGCTG	GAGC	CACTG	CCCAT	CGTGT	ACTAC	GTGGG	CCGCAAGCCCAAGG 741
Qy	1180	TGGAG	CAG	CTGT	CCAAC	ATGAT	CGTG	CGTT	CCTG	CAAGT	GCAG	CTGAGGCCCCGCCCCGC 1239
Db	742	TGGAG	CAG	CTGT	CCAAC	ATGAT	CGTG	CGTT	CCTG	CAAGT	GCAG	CTGAGGCCCCGCCCCGC 801
Qy	1240	CCACAG	CCCCG	CCCC	ACCCG	G	AGG	CCCCG	CCCC	ACCCCC	GCCCC	GCCCTCACCGGGG 1294
Db	802	CACN	CGGN	CNNNN	NGGN	MNGG	GGGG	GNCCCC	GGCG	CCGGG	CCCCG	CGGCCCGG 856

RESULT 8
 BX383773/c
 LOCUS BX383773 950 bp mRNA linear EST 08-MAY-2003
 DEFINITION BX383773 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DK001YA15 3-PRIME, mRNA sequence.
 ACCESSION BX383773
 VERSION BX383773.1 GI:30457168

Qy		597 TGGCTGTCCCTTTGATGTCAACCGAGTGTGTGCCGCAGTGCGTCACC CGCAGAGAGGCTATA	656
Db		703 TG GTT AT C T T T T G A T G T C A C C G G A G T T G T G C G G C A G T G G T T G A G C C G T G G A G G G G A A A T T	644
Qy		657 GAGGGTTTTTCGCCTCAGTGCCC ACTGTT CCTGTG ACAGCAA AGATAACA CA CT CCACGTG	716
Db		643 GA - GGCTTT GCCTTAGCGCCCCACTGCTCCTGTTACAGCAGGGATTACACACTGCAAGTG	585
Qy		717 GAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCC ACCATTAC CGGCATGAAC	776
Db		584 GACATCAACGGGTTTACTACCGGCCGCCGAGGTGACCTGGCCACCATT CATGGCATGAAC	525
Qy		777 CGGCCCTTCTCTGCTCCTCATGGCCACCCCGCTGGAGAGGG CCCAGCACCTGCACAGCTCC	836
Db		524 CGGCCTTCTCTGCTTCTCATGGCCACCCCGCTGGAGAGGG CCCAGCATCTGCAAAGCTCC	465
Qy		837 CGGCACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGAT	896
		:	
Db		464 CGGCACCGCCGA-----GCMCTGGAC	444
Qy		897 ACCA ACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGAC	956
		:	
Db		443 ACCA ACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTKAC	384
Qy		957 TTCCGGAAGGACCTGGGCTGGAAGTGGATT CATGAACCCAAGGGCTACCATGCCAAT TTC	1016
Db		383 TTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCGACT TC	324
Qy		1017 TGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCT	1076
Db		323 TGCCTCGGGCCCTGCCCCCTACATTGAGCCTGGACACGCAGTACAGCAAGGTCTCTGGCG	264
Qy		1077 CTGTACAACCAGCAACAACCCGGGCGCGTCGGCGGC GCCGTGCTGCGTGCCGCAGGCGCTG	1136
Db		263 CCTGTACAACCAGCATAAACCCGGGCGCCTCGGCGGC GCCGTG- T GCGTGCCGCAGGCGCTG	205
Qy		1137 GAGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAAC	1196
		:	
Db		204 GAGCCGCTGCCCATCGTGTACTACGTGGGCKKCAAGCCCAAGGTGGAGCAGCTGTCCNAC	145
Qy		1197 ATGATCGTGCGTTCTCTGCAAGTGCAGCTGAGGCCCGCCCCGCCCAAGCCCCGCCCCACC	1256
Db		144 ATGATCGTGCGCTCCTGCAAGTGCAGCTGAGATCCCGCCCCGCCCGCCCCGCCCGCGGAA	85
Qy		1257 CGGCAGGCCCCGGCCCCACCCCGCCCGC-----CTCACC GGGGCTGTATTTAAGGA	1307
Db		84 GGCCGCGGCCCAACCCCGCCCGCCCCSCTGCCTTGCCCATGGGGGCTGTATTTAAGNA	25
Qy		1308 -CATCGTGCCCCAAGCCCACTTGG	1330
Db		24 CCCCCCTGGCCCCAAGCCCACTTGG	1

```

cDNA clone CS0DD009YM06 5-PRIME, mRNA sequence.
ACCESSION      AL530081
VERSION        AL530081.2  GI:31067916
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 859)
AUTHORS        Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE          Full-length cDNA libraries and normalization
JOURNAL        Unpublished
COMMENT        On Feb 13, 2001 this sequence version replaced gi:12793574.
                Contact: Genoscope
                Genoscope - Centre National de Sequencage
                BP 191 91006 EVRY cedex - France
                Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
                was normalized. Library was constructed by Life Technologies, a
                division of Invitrogen. This sequence belongs to sequence cluster
                9160.r For more information about this cluster, see
                http://www.genoscope.cns.fr/
                cgi-bin/cluster.cgi?seq=CS0DD009BG03QP1&cluster=9160.r. Contact :
                Feng Liang Email : fliang@lifetech.com URL :
                http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
                Faraday Avenue Genoscope sequence ID : CS0DD009BG03QP1.

FEATURES
  source          Location/Qualifiers
                  1..859
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="CS0DD009YM06"
                  /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
                  /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
                  /note="1st strand cDNA was primed with a NotI-oligo(dT)
                  primer. Five prime end enriched, double-strand cDNA was
                  digested with Not I and cloned into the Not I and EcoR V
                  sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT      179 a      272 c      257 g      148 t      3 others
ORIGIN

Query Match          42.9%;  Score 580.6;  DB 9;  Length 859;
Best Local Similarity 87.3%;  Pred. No. 1.9e-116;
Matches 645;  Conservative 2;  Mismatches 91;  Indels 1;  Gaps 1;

Qy      10  CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 69
      ||  |||  ||||  ||  ||||||||||||  ||  ||||||||||  |||||||||||||
Db      121  CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTAC 180

Qy      70  TAGTGCTGACGCCTGGCCGGCCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGG 129
      |  |||||||||||||||||||||||||  ||||  ||||||||||  |||||||||||
Db      181  TGGTGCTGACGCCTGGCCGGCCGGCCGCCGACTATCCACCTGCAAGACTATCGACATGG 240

Qy      130  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 189
      |||||||||||||||||||||||||  ||||||||||  ||||||||||  ||||
Db      241  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 300

Qy      190  TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 249

```

Db	301	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	360
Qy	250	CTCTTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCAGC	309
Db	361	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCAGC	420
Qy	310	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC	369
Db	421	CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACG	480
Qy	370	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG	429
Db	481	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	540
Qy	430	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	489
Db	541	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	600
Qy	490	GGCTCAAGTTAAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCCT	549
Db	601	GGCTCAAGTTAAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCCT	660
Qy	550	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTG-GCTGTCCTTT	608
Db	661	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTGTTATCTTTT	720
Qy	609	GATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGC	668
		:	
Db	721	GATGTCACCGGAGTTGTGCGGCAGTKGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGC	780
Qy	669	CTCAGTGCCCACTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGG	728
		:	
Db	781	CTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGG	840
Qy	729	TTCAATTCTGGCCGCCGGG	747
Db	841	TTCACTACCGGCCGCCGAG	859

Db 301 AAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGT 360

```

Qy      499 TAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACC 558
        |||
Db      361 TAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGATACC 420

Qy      559 TCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTCACCG 618
        |||
Db      421 TCAGCAACCGGCTGCTGGCAGCCAGCGACTCGCCAGAGTGGTTATCTTTTGTATGTCACCG 480

Qy      619 GAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCC 678
        |||
Db      481 GAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCC 540

Qy      679 ACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTG 738
        |||
Db      541 ACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCG 600

Qy      739 GCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGG 798
        |||
Db      601 GCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTCATGG 660

Qy      799 CCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC 851
        |||
Db      661 CCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCGAGCC 713

```

RESULT 11

BE260971

LOCUS BE260971 717 bp mRNA linear EST 26-OCT-2000

DEFINITION 601153715F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3509931 5', mRNA sequence.

ACCESSION BE260971

VERSION BE260971.1 GI:9132709

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 717)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM193 row: m column: 04

High quality sequence stop: 713.

FEATURES Location/Qualifiers

source 1..717

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3509931"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_19"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 158 a 226 c 213 g 120 t
ORIGIN

Query Match 41.8%; Score 565.8; DB 10; Length 717;
Best Local Similarity 87.7%; Pred. No. 3.1e-113;
Matches 629; Conservative 0; Mismatches 87; Indels 1; Gaps 1;

```
Qy      140 GCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCC 199
          |||
Db       1  GCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCC 60

Qy      200 CCCGAGCCAGGGGGACGTGCCGCCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAA 259
          |||
Db       61 CCCGAGCCAGG-GGAGGTGCCGCCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTGTACAA 119

Qy      260 CAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAGGCGGA 319
          |||
Db      120 CAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCCGA 179

Qy      320 CTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATCTATGA 379
          |||
Db      180 CTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACGAAATCTATGA 239

Qy      380 TAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGA 439
          |||
Db      240 CAAGTTCAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGA 299

Qy      440 AGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTT 499
          |||
Db      300 AGCGGTACCTGAACCCGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTT 359

Qy      500 AAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCT 559
          |||
Db      360 AAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGATACCT 419

Qy      560 CAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGATGTCACCGG 619
          |||
Db      420 CAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTTATCTTTTGATGTCACCGG 479

Qy      620 AGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGTGCCCA 679
          |||
Db      480 AGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCA 539

Qy      680 CTGTTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGG 739
          |||
```

```

Db      540 CTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGG 599
Qy      740 CCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGC 799
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      600 CCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTTCCTGCTTCTCATGGC 659
Qy      800 CACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGA 856
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      660 CACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCGAGCCCTGGGA 716

```

RESULT 12

BQ675698

LOCUS BQ675698 902 bp mRNA linear EST 15-JUL-2002

DEFINITION AGENCOURT_8036532 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6211917
5', mRNA sequence.

ACCESSION BQ675698

VERSION BQ675698.1 GI:21786532

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 902)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2375 row: c column: 22

High quality sequence stop: 599.

FEATURES Location/Qualifiers

source 1..902

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6211917"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_102"

/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 194 a 292 c 255 g 161 t

ORIGIN

Query Match 41.7%; Score 564.8; DB 13; Length 902;
 Best Local Similarity 87.6%; Pred. No. 5.5e-113;
 Matches 617; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```

Qy      160 TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 219
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1   TCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCCCCGAGCCAGGGGGAGGTGC 60

Qy      220 CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 279
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61  CGCCCGGCCCGCTGCCCCGAGGCCGTGCTCGCCCTGTACAACAGCACCCGCGACCGGGTGG 120

Qy      280 CCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCA 339
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     121  CCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCCGACTACTACGCCAAGGAGGTCA 180

Qy      340 CCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAATTCAAGGGCACCCCCC 399
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     181  CCCGCGTGCTAATGGTGGAAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACAC 240

Qy      400 ACAGCTTATATATGCTGTTCAACACGTCCGAGCTCCGGAAGCGGTGCCGGAACCTGTAT 459
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     241  ACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGT 300

Qy      460 TGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG 519
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     301  TGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG 360

Qy      520 AGCTATACCAGAAATACAGCAATGATTCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCC 579
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     361  AGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCTGGCAC 420

Qy      580 CCAGTGACTCACCGAGTGGCTGTCTTTGATGTACCCGAGTTGTGCGGCAGTGGCTGA 639
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     421  CCAGCGACTCGCCAGAGTGGTTATCTTTTGTATGTACCCGAGTTGTGCGGCAGTGGTTGA 480

Qy      640 CCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTGTTCTCTGTGACAGCAAAG 699
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     481  GCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGG 540

Qy      700 ATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCA 759
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     541  ATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCA 600

Qy      760 CCATTACCGGCATGAACCGGCCCTTCTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCC 819
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     601  CCATTTCATGGCATGAACCGGCCCTTCTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCC 660

Qy      820 AGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAAC 863
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     661  AGCATCTGCCAAGCTCCCGGCACCGCCGAGCCCTGGGACACCAC 704
  
```

RESULT 13

BI818841

LOCUS BI818841 925 bp mRNA linear EST 04-OCT-2001

```

DEFINITION 603037307F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5178433 5',
            mRNA sequence.
ACCESSION  BI818841
VERSION    BI818841.1  GI:15930391
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

```

```
REFERENCE      1 (bases 1 to 925)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11444 row: n column: 02
High quality sequence stop: 874.

```

FEATURES             Location/Qualifiers
     source            1. .925
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:5178433"
                        /lab_host="DH10B"
                        /clone_lib="NIH_MGC_115"
                        /note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."

```

BASE COUNT	170 a	316 c	278 g	161 t
ORIGIN				

Query Match 39.7%; Score 537.6; DB 12; Length 925;
Best Local Similarity 87.5%; Pred. No. 4.8e-107;
Matches 611; Conservative 0; Mismatches 84; Indels 3; Gaps 2;

[illegible]


```

      |||
Db      483 ACCCCGCTGGAGAGGGCCAGCTTCTGCAAAGCTCCCGCACCGCCGA----- 436
Qy      861 AACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACG 920
      |||
Db      435 -----GCCCTGGACACCAACTATTGCTTCAGCTCCACG 403
Qy      921 GAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAG 980
      |||
Db      402 GAGAAGAACTGCTGCGTGCGGCAGCTGTACATTAAGTCCGCAAGGACCTCGGCTGGGAG 343
Qy      981 TGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATC 1040
      |||
Db      342 TGGATCCACGAGCCCAAGGGCTWCCATGCCAAGTCTGCCTCGGGCCCTGCCCCTACATT 283
Qy      1041 TGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAACCCGGGC 1100
      |||
Db      282 TGGAGCCTGGACACGACGAGGACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCCGG-- 225
Qy      1101 GCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTAC 1160
      |||
Db      224 GCCTCGGCGGCGCCGTGC-GCGTGCCGCAGGCGCTGGAGCCGCTGCCCTTCGTGTACTAC 166
Qy      1161 GTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAAGTGC 1220
      |||
Db      165 GTGGGCGCAAGCCCAAGGTGGGGCGGCTGTCCAACATGGTTCGTGCGCTCCTGCAAGGGC 106
Qy      1221 AGCTGAGGCCCCGCCCCGCCCCAGCCCCGCCCCAGGCCCCGCCCCACCCCCGC 1280
      |||
Db      105 CGCTGGGGTCCCCCGCCCCGCCCCGCCCCGCGCAGGCCCCGCCCCCGCCCCGCC 46
Qy      1281 CCGCCT 1286
      ||: ||
Db      45 CCKGCT 40

```

RESULT 15

BI084718/c

LOCUS BI084718 956 bp mRNA linear EST 20-JUN-2001

DEFINITION 602869722T1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5014421 3', mRNA sequence.

ACCESSION BI084718

VERSION BI084718.1 GI:14503048

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 956)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1821 row: 1 column: 06
 High quality sequence start: 3
 High quality sequence stop: 793.

FEATURES Location/Qualifiers
 source 1. .956
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5014421"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_102"
 /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-cDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 172 a 273 c 312 g 199 t
 ORIGIN

Query Match 39.7%; Score 536.6; DB 12; Length 956;
 Best Local Similarity 78.0%; Pred. No. 8e-107;
 Matches 757; Conservative 0; Mismatches 159; Indels 55; Gaps 7;

```

Qy      333 GAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAATTCAAGGGC 392
          ||||| | |||      ||||| ||||| | | ||| ||      | |||| |
Db      955 GAGGTTAACCGGTGCTAAATGGTGGAAACCCACCACGAATCTATGACCAGGTTTCAGCAGA 896

Qy      393 ACCCCCCACAGCTTATATATGCTGTTCAACACGTC--GGAGCTCCGGGAAGCGGTGCCGG 450
          ||      | || ||||| | ||||      || |      |||| ||| |||| ||
Db      895 GTACACACAGCCATAAATATGGTCTTTCACACATTTCAGGAGCTTCCGAGAACCGGTACCCT 836

Qy      451 AACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTG-----CTGAGGCTCAAGTTAAAAGT 505
          | | || |||      || | || || || || || || || || || || || || || ||
Db      835 GAACCGTGTTGGCTCTTCCCGGGGAAGAGCTGCCGTCTGCCTGAGGCTCAAGTTAAAAGT 776

Qy      506 -GGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCA 564
          ||| ||||| ||||| ||| |      |      |      | ||||| ||||| |||||
Db      775 GGGACCAGCACGTGGAGCTGTAACCAGGAATACAGCACAATTCTGGCGATACCTCAGCA 716

Qy      565 ACCGGCTGCTGGCCCCCAGTGA--CTCACCGGAGTGGCTGTCCTTTGATGTACCCGGAGT 622
          ||||| |||| | |||| || || || || || || || || || || || || || || ||
Db      715 ACCGGCTGGTGGCACCCAGCGAACTCGCCAGAGTGGTTATCTTTTGTATGTACCCGGAGT 656

Qy      623 TGTGCGGCAGTGGCTGACCCGAG-AGAGGCTATAGAGGGTTTTCGCCTCAGTGCCCACT 681
          ||||| ||||| |||| ||| || || || || || || || || || || || || || ||
Db      655 TGTGCGGCAGTGGTTGAGCCGTGGAAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCACT 596

Qy      682 GTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCC 741
          | ||||| ||||| ||||| ||||| || || || || || || || || || || || || ||

```

```

Db      595 GCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCC 536
Qy      742 GCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCCA 801
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      535 GCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTTCCTGCTTCTCATGGCCA 476
Qy      802 CCCCCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCA 861
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      475 CCCCCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC---- 420
Qy      862 ACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACGG 921
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      419 -----ACCAACTATTGCTTCAGCTCCACGG 395
Qy      922 AGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGT 981
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      394 AGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGT 335
Qy      982 GGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCT 1041
      ||||| || || || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      334 GGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTACATTT 275
Qy      1042 GGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCG 1101
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      274 GGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCGGGCG 215
Qy      1102 CGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACG 1161
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      214 CCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACG 155
Qy      1162 TGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCA 1221
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      154 TGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAAGTGCA 95
Qy      1222 GCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGCCCCACCCCGGCC 1281
      ||||| ||||| ||||| || || || ||||| ||||| ||||| ||||| |||||
Db      94 GCTGAGGTCCCGCCCCGCC-----CCGCCCCGCCCGGCAGGCCCGGCCCCACCCCGGCC 40
Qy      1282 CGCCTCACCGG 1292
      ||||| || || ||
Db      39 CGCCCCCGCTG 29

```

Search completed: October 28, 2003, 00:08:25
Job time : 3362.54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2003, 15:56:23 ; Search time 5244 Seconds
(without alignments)
10617.468 Million cell updates/sec

Title: US-10-017-372E-36

Perfect score: 1361
Sequence: 1 tggaccagatggcgctt.....cgattaaagcgccgcgact 1361

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1221.2	89.7	3206	4	PIGTGFB1A M23703 Sus scrofa
2	1211.8	89.0	1326	6	AX338213 Sequence
3	1201.8	88.3	1750	4	GCTGFB1 X12373 Porcine mRN
4	1131.6	83.1	1605	4	SSTGFB1 Y00111 Porcine mRN
5	1110.2	81.6	2221	4	AF461808 Sus scrofa
6	990.6	72.8	1369	4	DOGTFB1A L34956 Canine tran
7	979.6	72.0	2527	6	E00973 E00973 cDNA encodi
8	978.8	71.9	2537	6	A06669 A06669 Synthetic m
9	972.8	71.5	1173	4	ATGFB1 X76916 O.aries mRN
10	972.4	71.4	1780	9	BC000125 Homo sapi
11	972.4	71.4	1780	9	BC001180 Homo sapi
12	971	71.3	1561	9	AGMTGFB M16658 Simian tran
13	968.2	71.1	1821	6	E03028 E03028 DNA encodin
14	967.8	71.1	1746	9	BC022242 Homo sapi
15	966.6	71.0	1560	6	I06216 I06216 Sequence 2
16	965.8	71.0	2745	9	HSTGFB1 X02812 Human mRNA
17	964.6	70.9	1560	6	I08268 I08268 Sequence 2
18	947.2	69.6	1569	6	I08221 I08221 Sequence 3
19	942.2	69.2	1561	6	I08275 I08275 Sequence 3
20	927.6	68.2	1173	12	BT007245 Homo sapi
21	927.6	68.2	1173	12	BT007866 Synthetic
22	918.4	67.5	1173	4	ECRGFB1 X99438 E.caballus
23	918.2	67.5	1176	6	AX615127 Sequence
24	916.6	67.3	1176	6	AX481432 Sequence
25	916.6	67.3	1176	6	AX615128 Sequence
26	912	67.0	1187	4	AF175709 Equus cab
27	889.2	65.3	1597	10	AF191297 Cavia por
28	860.8	63.2	1641	10	BC013738 Mus muscu
29	847	62.2	1579	10	MUSTGFRNA M13177 Mouse trans
30	847	62.2	2094	10	MNU009862 AJ009862 Mus muscu
31	827.4	60.8	1585	10	RNTGFB1 X52498 Rat mRNA fo
32	815.6	59.9	1117	4	BOVTGFB M36271 Bovine tran
33	791	58.1	1125	10	AF480858 Sigmodon
34	682.2	50.1	1675	6	I03310 I03310 Sequence 1
35	654	48.1	1376	6	AX528533 Sequence
36	652.4	47.9	1389	6	AX528619 Sequence
37	591.4	43.5	1352	6	AX528535 Sequence
38	586.6	43.1	1350	6	AX528615 Sequence
39	543.6	39.9	699	6	I05434 I05434 Sequence 4
40	357.6	26.3	650	6	AX336646 Sequence
41	357.6	26.3	650	9	HDMTGFBA M38449 Homo sapien
42	357.6	26.3	862	6	I03312 I03312 Sequence 3
43	354.4	26.0	489	6	AX455100 Sequence
44	327.2	24.0	469	10	MATGFB1 X60296 M.auratus m
45	318.2	23.4	1256	5	CHKTGFB4 M31160 Gallus gall

ALIGNMENTS

RESULT 1	PIGTGFB1A	Sus scrofa	transforming growth factor beta-1 mRNA	linear	MAM 31-MAR-1995
LOCUS	M23703				
DEFINITION	M23703				
ACCESSION	M23703.1	GI:755044			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					

TITLE Inducible plasmid vector encoding tgf₂(b) and uses thereof
 JOURNAL Patent: WO 0181404-A 1 01-NOV-2001;
 THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

FEATURES
 source Location/Qualifiers
 1..1326
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 16..1188
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD20538.1"
 /db_xref="GI:18128751"
 /translation="MAPSGRLRLPLLLPLLLVLTGPRPAAGSLSTCKTIDMELVKRK
 REIAEQILSLKSLSPSGDVPGLPEAVLALNSYDRVRVAGSEVEPEPEAD
 YVAKETRVLMVSGNOIYDKFGTSPHLYMLFNTSELREAVPEPVLLSRAELRLRL
 KLVKEOHVELYOKYSNDWRLSNRLAPSDPEMLSDVTGVVROWLTTRREAIGFR
 LSAHSSSDSKDNTLHWEINFGNSGRGLDATTGHMNRPLFLMLMATPLERAQHLSSRH
 RRALDNYCESSTKKNCCVRLQYIDFRDLGKHWIHEPKGYHANFCLGPCPYINSLDT
 QYSKVLALYNQHNHPGASAPCCVPQALEPLPIVTVGRKPRVQLSNMIVRSKUS"
 BASE COUNT 263 a 438 c 392 g 233 t
 ORIGIN

Query Match 89.0%; Score 1211.8; DB 6; Length 1326;
 Best Local Similarity 95.88; Pred. No. 1.3e-209;
 Matches 1299; Conservative 0; Mismatches 7; Indels 53; Gaps 3;

QY 1 TGTACCGAGATGGCGCTTTCGGGGCTGCGGGCTTTGGCGCTGCTGCTCCCGCTGCTGTG 60
 DB 6 TGTACCGAGATGGCGCTTTCGGGGCTGCGGGCTTTGGCGCTGCTGCTCCCGCTGCTGTG 65

QY 61 GCTGCTAGTCTACGCTTGGCGCGCGCGCGCGCGGACTGTCCACCTGCAAGACCATCGA 120
 DB 66 GCTGCTAGTCTACGCTTGGCGCGCGCGCGCGCGGACTGTCCACCTGCAAGACCATCGA 125

QY 121 CATGGAGCTGGTGAAGCGGAGCGCATCGAGGCGCATTCGCGCGCAGATTCTGTCCAAGCT 180
 DB 126 CATGGAGCTGGTGAAGCGGAGCGCATCGAGGCGCATTCGCGCGCAGATTCTGTCCAAGCT 185

QY 181 TCGGCTTGGCAGCG 240
 DB 186 TCGGCTTGGCAGCG 245

QY 241 ACTGGCTCTTTACACAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 DB 246 ACTGGCTCTTTACACAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 305

QY 301 CGAGCCAGCGCGGACTACTACGCCAAGGAGGTACCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 DB 306 CGAGCCAGCGCGGACTACTACGCCAAGGAGGTACCGCGCGCGCGCGCGCGCGCGCGCGCG 365

QY 361 CAACCAATCTATGATAAAATCAAGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 DB 366 CAACCAATCTATGATAAAATCAAGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 425

QY 421 GTCGAGCTCCGGAAGCGGTGCGGAACTGTTATGCTCTCTCGGCGCAGAGTGGCGCT 480
 DB 426 GTCGAGCTCCGGAAGCGGTGCGGAACTGTTATGCTCTCTCGGCGCAGAGTGGCGCT 485

QY 481 GCTGAGGCTCAAGTAAAGCTGAGGAGCAGCTGGAGCTATACCAAAATACAGCAATCA 540
 DB 486 GCTGAGGCTCAAGTAAAGCTGAGGAGCAGCTGGAGCTATACCAAAATACAGCAATCA 545

QY 541 TTCCTGGCGCTACCTACGACACCGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 DB 546 TTCCTGGCGCTACCTACGACACCGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 605

QY 601 CTTTGTATGTCACCGGAGTCTGCGGCGAGTGGCTGACCGCGCAGAGAGGCTATAGAGGTTT 660
 DB 606 CTTTGTATGTCACCGGAGTCTGCGGCGAGTGGCTGACCGCGCAGAGAGGCTATAGAGGTTT 665

QY 661 TCGGCTCAGTGGCGGCTCTTCTCTGACACAGAGATACACACTCCAGCTGGAATTTAA 720

Db 666 TCGCCICAGTGCACACTCTTCTCTGACACAGAAAGATAACACACTCCACGTGGAATTA 725

QY 721 CGGTTCAATTTCTGGCGCGCGGGTCACTTGGCCACCATTCAGCGCATGAACCGGCCCTT 780

Db 726 CGGTTCAATTTCTGGCGCGCGGGTCACTTGGCCACCATTCAGCGCATGAACCGGCCCTT 785

QY 781 CTGCTCTCATGGCCACCGCGCTGGAGAGGGCGCCAGCACTGCAAGCTCCCGGACCG 840

Db 786 CTGCTCTCATGGCCACCGCGCTGGAGAGGGCGCCAGCACTGCAAGCTCCCGGACCG 845

QY 841 CCGAGCCCTGGATACCAACAGCTACCCATAGACGTGCCAGACTACGCATCTCTGGCCCT 900

Db 846 CCGA-----GCGCT 854

QY 901 GGATACCAACTACTGCTTTCAGCTCCACGGAGAAAGAACTGCTGGTGGCGCAGCTCTACAT 960

Db 955 GGATACCAACTACTGCTTTCAGCTCCACGGAGAAAGAACTGCTGGTGGCGCAGCTCTACAT 914

QY 961 TGACTTCGGAAGGACCTGGCTGGAGTGGATTGATGAACCAAGGGCTACCATGCCAA 1020

Db 915 TGACTTCGGAAGGACCTGGCTGGAGTGGATTGATGAACCAAGGGCTACCATGCCAA 974

QY 1021 TTTCTGCTGGCGGCTTCTCCCTACATCTGAGCCTAGACACTCAGTACAGCAAGTCTCT 1080

Db 975 TTTCTGCTGGCGGCTTCTCCCTACATCTGAGCCTAGACACTCAGTACAGCAAGTCTCT 1034

QY 1081 GGCTCTGTACAAACAGACAAACCGCGGCGTGGCGGGCGCGTGGTGGCGGAGCG 1140

Db 1035 GGCTCTGTACAAACAGACAAACCGCGGCGTGGCGGGCGCGTGGTGGCGGAGCG 1094

QY 1141 GCTGAGCCACTGCGCATCGTGTACTGAGTGGCGCGCAAGCCCAAGSTGGAGCAGCTGTC 1200

Db 1095 GCTGAGCCACTGCGCATCGTGTACTGAGTGGCGCGCAAGCCCAAGSTGGAGCAGCTGTC 1154

QY 1201 CAACATGATCGTGGTTCCTGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1260

Db 1155 CAACATGATCGTGGTTCCTGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1213

QY 1261 CACCGCGCAGCG 1320

Db 1214 CACCGCGCAGCG 1273

QY 1321 TGCCCCAAGCGCGCTTGGGATCGGATTAAAGCGCGCGCGA 1359

Db 1274 TGCCCCAAGCGCGC-TGGGATCGGATTAAAGGTGAGAGA 1311

RESULT 3
 GGTGFB1
 LOCUS Porcine mRNA for transforming growth factor-beta 1. 1750 bp mRNA linear MAM 27-MAR-1996

DEFINITION
 ACCESSION X12373
 VERSION X12373.1 GI:63808
 KEYWORDS Transforming growth factor-beta 1.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa

REFERENCE
 1 (bases 1 to 1750)
 AUTHORS Jakowlew S.B., Dillard P.J., Sporn M.B. and Roberts A.B.
 TITLE Nucleotide sequence of chicken transforming growth factor-beta 1 (TGF-beta 1)
 JOURNAL Nucleic Acids Res. 16 (17), 8730 (1988)
 MEDLINE 8835639
 PUBMED 3166520

REFERENCE
 2 (bases 1 to 1750)
 AUTHORS Jakowlew S.B.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-1988) Jakowlew S.B., National Institute of Health, National Cancer Institute, Laboratory of Chemoprevention, Building 41, Room B902, Bethesda, Maryland 20892, USA

COMMENT
 The submitters believe that the chicken cDNA library was contaminated with porcine cDNA, and that the sequence is infect

	Best Local Similarity	94.6%	Indels	54	Gaps	2
	Matches	1286	Conservative	0	Mismatches	17
Qy	6	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGGCGCTGCTGCTGCGCGTGTGTGGTGTGC	65			
Db	442	CCCCATCCGCGCTTCGGGGCTTGGGCTCTTGGCGCTGCTGCTGCGCGTGTGTGGTGTGC	501			
Qy	66	TAGTCTCAGCGCTGGCGCGCGCGCGGAGCTGTCCACCTGCAGACCATCGACATGG	125			
Db	502	TAGTCTCAGCGCTGGCGCGCGCGCGGAGCTGTCCACCTGCAGACCATCGACATGG	561			
Qy	126	AGCTGGTGAAGCGGAAGCGCATTCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC	185			
Db	562	AGCTGGTGAAGCGGAAGCGCATTCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC	621			
Qy	186	TTGCGACGCCCCCGAGCGCAGGGAGCGTGGCGCCCGCGCGCGCTGCTGAGGCGGTACTGG	245			
Db	622	TGCGACGCCCCCGAGCGCAGGGAGCGTGGCGCCCGCGCGCTGCTGAGGCGGTACTGG	681			
Qy	246	CTCTTTTACAACAGTACCCGCGACCGGGTAGCGGGGAAAAGTGTGAAACGGAGCCCGAGC	305			
Db	682	CTCTTTTACAACAGTACCCGCGACCGGGTAGCGGGGAAAAGTGTGAAACGGAGCCCGAGC	741			
Qy	306	CAGAGCGGACTACTACGCCAAGGAGGTACCCGCGTGTAAATGTGTGGAACGGCAACC	365			
Db	742	CAGAGCGGACTACTACGCCAAGGAGGTACCCGCGTGTAAATGTGTGGAACGGCAACC	801			
Qy	366	AAATCTATGATAAATTCAAAGGCGACCCCGCACAGCTTATATATGCTGTTCACACAGCTCGG	425			
Db	802	AAATCTATGATAAATTCAAAGGCGACCCCGCACAGCTTATATATGCTGTTCACACAGCTCGG	861			
Qy	426	AGTCCGGGAAGCGGTGCCGGAACCTGATTGTCTCTCGGGCAGAGCTGCCCTGCTGA	485			
Db	862	AGTCCGGGAAGCGGTGCCGGAACCTGATTGTCTCTCGGGCAGAGCTGCCCTGCTGA	921			
Qy	486	GGCTCAAGTTAAAGTGGAGCAGCAGCTGGAGCTATACCAAGAATACACCAATGATTCCT	545			
Db	922	GGCTCAAGTTAAAGTGGAGCAGCAGCTGGAGCTATACCAAGAATACACCAATGATTCCT	981			

gene	/chromosome="6"	
	1. .2221	
CDS	/gene="TGFB1"	
	1. .1173	
	/gene="TGFB1"	
	/note="cytokine"	
	/codon_start=1	
	/product="transforming growth factor beta 1"	
	/protein_id="AAL57902.1"	
	/db_xref="GI:18042251"	
	/translation="MPSPGLRLPLLLPLLLVLPGRPAAGLSCTKIDMELVKKR RIEARGIILSKLRLASPPSGDVPVPGPLPEAVIALYNSTRDRVAGESVEPEPEAD YIAKEVTRVLVYQKISNDMSYLSNRLIAPSDSPWISFDVTVVROWLRRRAELRLR LKUVQSVELYQKISNDMSYLSNRLIAPSDSPWISFDVTVVROWLRRRAELRGR LSAHCSDKNTLHVEINFGNSRRGLDATHGMRPFLLLMATPLERAHLHSRR RRALDNTYCFSSYTKNCVQRLYIDFKDLGKWIHEPKGYHANFCLGPCPVILSLDT QYSKVLALYNQHNHNGASAAAPCCVQALPLPIVYVGRKPKVQLSNMIVRSCKS"	
3'UTR	1174. .2221	
	/gene="TGFB1"	
BASE COUNT	504 a	616 c 649 g 452 t
ORIGIN		
Query Match	81.6%	Score 1110.2; DB 4; Length 2221;
Best Local Similarity	95.2%	Pred. No. 3e-191;
Matches 1176; Conservative	0; Mismatches	8; Indels 51; Gaps 1;
QY	11	ATGGCGCTTCGGGGCTGCGGCTCTTCCGGTGTGCTGCGGCTGCTGGCTGTAGTG 70
Db	1	ATGGCGCTTCGGGGCTGCGGCTCTTCCGGTGTGCTGCGGCTGCTGGCTGTAGTG 60
QY	71	CTGACGCTTGGCGCGCGCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 130
Db	61	CTGACGCTTGGCGCGCGCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
QY	131	GTGAAGCGGAAGCGCATGCAAGCATTCGCGGCGCAGATTCTGTCCAAAGCTTCGGCTTGC 190
Db	121	GTGAAGCGGAGGCATCGAGCCATTTCGCGGCGCAGATTCTGTCCAAAGCTTCGGCTTGC 180
QY	191	AGCCCCCGAGCGAGGGGAGCTGCGCGCGCGCGCGCTGCTGAGGAGTACTGGCTCTT 250
Db	181	AGCCCCCGAGCGAGGGGAGCTGCGCGCGCGCGCGCTGCTGAGGAGTACTGGCTCTT 240
QY	251	TACAACAGTACCGCGACCGCGGTAGCGGGGAAAGTGTGAAACCGGAGCCCGAGCCAGAG 310
Db	241	TACAACAGTACCGCGACCGCGGTAGCGGGGAAAGTGTGAAACCGGAGCCCGAGCCAGAG 300
QY	311	CGGACTACTACGCCAAGGAGTTCACCGCGTGTAAATGGTGGAAAGCGCAACCAATC 370
Db	301	CGGACTACTACGCCAAGGAGTTCACCGCGTGTAAATGGTGGAAAGCGCAACCAATC 360
QY	371	TATGATAAATCAAGGGCACCCCGACAGCTTATATGCTGTTCACACGTCGGAGCTC 430
Db	361	TATGATAAATCAAGGGCACCCCGACAGCTTATATGCTGTTCACACGTCGGAGCTC 420
QY	431	CGGGAAGCGGTGCGGAACCTGTATTGCTCTCTCGGCGAGAGCTCGGCTGCTGAGGCTC 490
Db	421	CGGGAAGCGGTGCGGAACCTGTATTGCTCTCTCGGCGAGAGCTCGGCTGCTGAGGCTC 480
QY	491	AAGTTAAAGTGGAGCAGCAGCTGGAGCTATACCAAGAAATACAGCAATGATTCCTGGCGC 550
Db	481	AAGTTAAAGTGGAGCAGCAGCTGGAGCTATACCAAGAAATACAGCAATGATTCCTGGCGC 540
QY	551	TACCTCAGCAACCGGCTGCTGCGCGCGAGTACATCAGCGGAGTGGCTGCTTTGATGTC 610
Db	541	TACCTCAGCAACCGGCTGCTGCGCGCGAGTACATCAGCGGAGTGGCTGCTTTGATGTC 600
QY	611	ACCGGAGTGTGCGGCGAGTGGTGAACCGCAGAGGCTATAGAGGTTTTTCGCTCAGT 670
Db	601	ACCGGAGTGTGCGGCGAGTGGTGAACCGCAGAGGCTATAGAGGTTTTTCGCTCAGT 660
QY	671	GCCCACTCTTCTCTGACAGCAAGATAACACACTCCACGCTGGAATTAACGGGTTCAAT 730
Db	661	GCCCACTGTGTTCTGTCAGAGCAAGATAACACACTCCACGCTGGAATTAACGGGTTCAAT 720
QY	731	TCTGGCGCGGGGTGACCTGGCCACCATTCAGGCATGAACGGCCCTTCCTGCTCCTC 790
Db	721	TCTGGCGCGGGGTGACCTGGCCACCATTCAGGCATGAACGGCCCTTCCTGCTCCTC 780
QY	791	ATGGCCACCCCGCTGGAGAGGGCCAGCACTGCACAGTCCCGGCGACCGCGAGCCCTG 850
Db	781	ATGGCCACCCCGCTGGAGAGGGCCAGCACTGCACAGTCCCGGCGACCGCGCGA----- 834
QY	851	GATACCAACAGCTACCCATAGGAGCTGCCAGACTACGCATCTCTGGCCCTGGATACCAAC 910
Db	835	-----GCCCTGGATACCAAC 849
QY	911	TACTGCTTACGTCCAGGAGAGAACTGTGGTGGCGGAGCTCTACATTTGACTTCCGG 970
Db	850	TACTGCTTACGTCCAGGAGAGAACTGTGGTGGCGGAGCTCTACATTTGACTTCCGG 909
QY	971	AAGGACCTGGGCTGGAAGTGAATTCATGAACCCAGGGCTACCATGCCAATTTCTGCCTG 1030
Db	910	AAGGACCTGGGCTGGAAGTGAATTCATGAACCCAGGGCTACCATGCCAATTTCTGCCTG 969
QY	1031	GGGCCCTGCTCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGTCTCTGGCTCTGTAC 1090
Db	970	GGGCCCTGCTCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGTCTCTGGCTCTGTAC 1029
QY	1091	AACAGCAACACCCGGCGCGCTGCGGCGCGCGCTGCTGCGCGCAGAGCGCTGGAGCCA 1150
Db	1030	AACAGCAACACCCGGCGCGCTGCGGCGCGCGCTGCTGCGCGCAGAGCGCTGGAGCCA 1089
QY	1151	CTGCCCATCTGTACTAGCTGGGCGCGCAAGCTGAGGAGCTGTCCAAACATGATC 1210
Db	1090	CTGCCCATCTGTACTAGCTGGGCGCGCAAGCTGAGGAGCTGTCCAAACATGATC 1149
QY	1211	GTGCGTTCTCTCAAGTGCAGCTGAGGCGCGCGCC 1245
Db	1150	GTGCGTTCTCTCAAGTGCAGCTGAGGCGCGCGCAAGCC 1184
RESULT 6		
DOGTGFB1A		
LOCUS	Canine transforming growth factor-beta 1 (TGFB1) mRNA, complete cds.	1369 bp mRNA linear MAM 30-OCT-1994
DEFINITION		
ACCESSION	L34956	
VERSION	L34956.1 GI:516071	
KEYWORDS	homologue; transforming growth factor-beta 1.	
SOURCE	Canis familiaris (dog)	
ORGANISM	Canis familiaris	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.	
AUTHORS	Manning,A.M., Auchampach,J.A., Drong,R.F. and Slightom,J.L.	
TITLE	Cloning of a canine cDNA homologous to human transforming growth factor-beta 1 (TGFBeta1)	
JOURNAL	Unpublished (1994)	
COMMENT	Original source text: Canis familiaris adult jugular vein endothelial cDNA to mRNA.	
FEATURES	Location/Qualifiers	
source	1. .1369	
gene	/organism="Canis familiaris"	
5'UTR	/mol_type="mRNA"	
CDS	/db_xref="taxon:9615"	
	/cell_type="LPS-activated"	
	/tissue_type="jugular vein endothelial"	
	/dev_stage="adult"	
	1. .1369	
	/gene="TGFB1"	
	1. .57	
	/gene="TGFB1"	
	58. .1230	
	/gene="TGFB1"	
	/function="anti-inflammatory agent"	
	/note="precursor"	

```

/codon_start=1
/product="transforming growth factor-beta 1"
/protein_id="AA51458.1"
/db_xref="GI:516072"
/translation="MPPSGRLRLPLLRLVLPGRPAAGLSTCKTIDMELVKRK
RIEATRGQILSKLRSSPPSQGEVPPVLPPEALYNSTRDVRVAGEPEPEPAD
YIAKEVTRVLWNTNKIKYKVKSPHSIYMLFNTSELREAVPEPVLSSRAELRLRL
KLAEOHVELYOKYSNDNSRYLSNRLAPSDTPEMLSPDVTGVVWMLSHGVEGFR
LSAHCSDSKDNTLOVDINGFSRRGDLATIHGMNRPFLLMATPLERAOHLHSSRO
RRALDNYCFSS"EXKCCVROQLYIDFRDLGKHNHERPKGYHANFCLGCPYINSLDT
QYSKJALNQHNPASAPCCVPQALEPLPIYTVGNRPVQBSNMIVRSCKCS"
1039..1227
mat_peptide
/gene="TGFB1"
/product="transforming growth factor-beta 1"
1231..1369
/gene="TGFB1"
BASE COUNT 264 a 473 c 415 g 216 t 1 others
ORIGIN
Query Match 72.8% Score 990.6; DB 4; Length 1369;
Best Local Similarity 86.8%; Pred. No. 1.5e-169;
Matches 1148; Conservative 0; Mismatches 125; Indels 53; Gaps 3;
QY 6 CCGNATGGCGCTCGGGCTCGTCCGCTGCTGCGGCTGCTGCGGCTGCTGGCTGC 65
DB 53 CCCCCATGCGCGCTCGGGCTCGGCTGCTGCGGCTGCTGCGGCTGCTGGCTGC 112
QY 66 TAGTGTGAGCGCTGCGCGCGCGCGCGGAGTGTCCACCTGCAAGACATCGACATGG 125
DB 113 TAGTGTGAGCGCTGCGCGCGCGCGCGGAGTGTCCACCTGCAAGACATCGACATGG 172
QY 126 AGCTGGTGAAGCGAAGCGCATGAGGCCATTGCGCGCCAGATTCTGTCCAAGCTTCGGC 185
DB 173 AACTGTGAAGCGAAGCGCATGAGGCCATCGCGCGCCAGATTCTGTCCAAGCTTCGGC 232
QY 186 TTGCGACCCCGGAGCGACGGGAGTGCGCGCGCGCGCGCTGCTGAGGCACTGACTGG 245
DB 233 TCTCAGCGCCCGGAGCGACGGGAGTGCGCGCGCGCGCTGCGCGGAGCGGCTGCTGG 292
QY 246 CTCTTTACACACTACCGCGACCGGTAGCGGGGAAAGTGTGCAACCGGAGCGCGAGC 305
DB 293 CCCTCTACACAGACCGCGGCGGCTGCGGGGAGAGCGCGGAGCGCGGAGCGCGAGC 352
QY 306 CAGAGCGGAGTACTAGCCCAAGGAGTCAACCGCGTGTCTAATGTGGAAAGCGGCAACC 365
DB 353 CCGAGCGGAGTACTAGCCCAAGGAGTCAACCGCGTGTCTAATGTGGAAAGCGGCAACC 412
QY 366 AATCTATGATAAATTCAGGGGACCGCCCGACAGCTTATATGCTGTTCAACACGTGG 425
DB 413 AATCTATGAGAAGTCAAGAAAGTCCGACAGCATATATGCTCTTCAACACATCAG 472
QY 426 AGCTCCGGGAGGCGTCCGCAACCTGTATTGCTCTCGGGCAGAGCTCGCGCTGCTCA 485
DB 473 AGCTCCGAGAGCAGTCCGAGCGCTGTGCTCTCGGGCAGAGTTCGCGCTGCTGA 532
QY 486 GGCTCAAGTTAAAGTGGACAGCAGTGGAGTATACAGAAATACAGCAATGATTCCT 545
DB 533 GGCTCAAGTTAAAGCGGAGCAGCATGTGGAGTGTACCAAGAAATATAGCAATGATTCCT 592
QY 546 GGCCTACCTCAGCAACCGGCTGCTGGCGCCAGTGACTCACCGGAGTGCTGCTTTG 605
DB 593 GGCCTACCTCAGCAACCGGCTGCTGGCGCCAGCGACGCGCAATGCTGCTTTG 652
QY 606 ATGTCACGGAGTGTGGCGAGTGGCTGACCGCGAGAGAGGCTATAGAGGGTTTTTCGCC 665
DB 653 ATGTCAGTGGAGTGTGGAGGAGTGGCTGAGCCATGGAGGGAAGTCGAGGGCTTTTCGCC 712
QY 666 TCAGTGCCACTCTCTCTGACAGCAAGATTAACACACTCCACGTGGAAATTAACGGGT 725
DB 713 TCAGTGCCACTCTCTCTGACAGCAAGATTAACACACTCCACGTGGAAATTAACGGGT 772
QY 726 TCAATTTCTGCGCGCGGGGTGACCTGGCCACCATTTACGGCATGAACCGGCTTCTCTGC 785

```

```

773 TCAGTTCAGCGCGCGGAGTGACCTGGCCACCATTCATGGCATGAACCGACCTTCCTGC 832
786 TCCTCATGCCACCGCGCTGGAGAGGGCCAGCACTGCACAGCTCCCGGCGACCGCCGAG 845
833 TCCTCATGCCACCGCGCTGGAGAGGGCCAGCACTGCACAGCTCCCGGCGACCGCG-- 890
846 CCCTGGATACCAACAGTACCCCATAGGAGTGCAGACTACGCATCTCTGGCGCCTGGATA 905
891 -----GGCCCTGGACA 901
906 CCAACTACTGCTTCACTGCACGGAGAAACTGCTGCGTGGGCGAGCTCTACATTGACT 965
902 CCAACTACTGCTTCACTGCACGGAGAAACTGCTGCGTCCGCGAGCTCTACATTGACT 961
966 TCGGGAAGGACCTGGGCTGGAAGTGATTCTCAACCCCAAGGCTTACCATGCCAATTTCT 1025
962 TCGCAAGGATCTGGGCTGGAAGTGATCCATGAGGCCAAGGTTACCAAGCTTACTTCT 1021
1026 GCTGGGGCCCTGTCCCTACATCTGAGGCTAGACACTCAGTACAGCAAGTCTCTGGGCTC 1085
1022 GCCTGGGGCCCTGCTTACATTTGGAGCTGGACACGAGTACAGCAAGTCTCTGGGCC 1081
1086 TGTACACACAGCACACCGCGGCGCTGCGCGCGCGCTGCTGCGTGGCGGAGCGCTGG 1145
1082 TGTACACACAGCACACCGCGGCGCTGCGCGCGCGCTGCTGCGTGGCGGAGCGCTGG 1141
1146 AGCACTGCCCATCTGTACTAGTGGCGCGCAAGCCCAAGTGGAGCAGCTGTCCAACA 1205
1142 AGCACTGCCCATCTGTACTAGTGGCGCGCAAGCCCAAGTGGAGCAGCTGTCCAACA 1201
1206 TGTCTGCTTCTCTCAAGTGCAGCTGAGGCGCGCGCGCG--CCACAGCGCGCGCCACC 1264
1202 TGATCTGTCGCTCTGCAAGTGCAGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCC 1261
1265 CGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1324
1262 CGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
1325 CCAAGC 1330
1321 CCAGGC 1326
RESULTS 7
LOCUS E00973 2527 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding human TGF-beta.
ACCESSION E00973
VERSION E00973.1 GI:2169234
KEYWORDS JP 1986219395-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2527)
AUTHORS Riku,M.A.D.D. and Debitsudo,B.G.
TITLE NUCLEIC ACID CODED WITH TGF-BETA AND ITS USE
PATENT: JP 1986219395-A 1 29-SEP-1986;
JOURNAL GENETECH INC
COMMENT
OS human
PN JP 1986219395-A/1
PD 29-SEP-1986
PF 20-MAR-1986 JP 1986064661
PR 22-MAR-1985 US 85 715142
PI RIKU MAIKERU ANDORE DERINKU, DEBITSUDO BANNOOMAN GETSUDEBU PC
C12P21/00,C12N1/00,C12N5/00,C12Q1/68,(C12P21/00, PC
C12R1-91),
PC (C12N1/00,C12R1:19),(C12N5/00,C12R1:91),(C12N15/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=placenta and glyoblastoma; FH Key

```

[illegible]

```
BASE COUNT      473 a   893 c   739 g   432 t
ORIGIN
RRALDINNYCFSSSTERNKCCVRLQYIDFRKDLGKWKIHERPKYHANFCLGCPYTWSLDT
OYSKVILALYNHNPFGASAPCCVPQALEPLPIVYVGRKPKVQLSNMIVRSCKCS"

Query Match      71.9%; Score 978.8; DB 6; Length 2537;
Best Local Similarity 84.8%; Pred. No. 1.9e-167;
Matches 1146; Conservative 0; Mismatches 147; Indels 59; Gaps 2;

QY      6  CCGNATGCGCCCTCGGGGCTCGGCTCTTCCGCTGCTGCTGCGGCTGCTGTGGCTGC 65
DB      837  CCCCCATCGCCCTCGGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAC 896
QY      66  TAGTGCTGACGCTTGGCCGCGCCGCGCGGAGCTGTCCACCTGCAAGACATCGACATGG 125
DB      897  TGTGCTGACGCTTGGCCGCGCCGCGGAGCTATCCACCTGCAAGACTATCGACATGG 956
QY      126  AGCTGTTGAAGCGGAAGCGATGAGGCGCATTCGCGGCGAGATCTGTCCAGCTTCGCGC 185
DB      957  AGCTGTTGAAGCGGAAGCGATGAGGCGCATTCGCGGCGAGATCTGTCCAGCTTCGCGC 1016
QY      186  TTGCGACGCCCGGAGCAAGGAGCTGCGCGCCGCGCGGCTGCTGAGGCGACTGCTGG 245
DB      1017  TCGCCAGCGCCCGGAGCAAGGAGCTGCGCGCCGCGCGGCTGCGCGGCGGCTGCTGCG 1076
QY      246  CTCCTTTACAACAGTACCGGCGACCGGTAGCGGGGAAAGTGTGCAACCGGAGCGCGAGC 305
DB      1077  CCCTGTGTAACAGCACCGCGGAGCGGGGTGCGGGGAGAGTGCAGAACCGGAGCGCGAGC 1136
QY      306  CAGAGCGGAGCTACTACGCAAGGAGTGCACCGCGTCTGATGTTGGAAGCGGCAACG 365
DB      1137  CTGAGCGCGAGCTACTACGCAAGGAGTGCACCGCGTCTGATGTTGGAAGCGGCAACG 1196
QY      366  AAATCTATGATAAATCAAGGGCACCGCCCGACAGCTTATATGCTGTTCACACAGCTCGG 425
DB      1197  AAATCTATGATAAATCAAGGGCACAGTACACAGCATATATATGCTTTCACACATCAG 1256
QY      426  AGCTCCGGGAGCGGTGCGGGAACCTGTTATGCTCTCGGGGAGAGTGCAGAACCGGAGC 485
DB      1257  AGCTCCGGGAGCGGTGCGGGAACCTGTTATGCTCTCGGGGAGAGTGCAGAACCGGAGC 1316
QY      486  GGCTCAAGTTAAAGTGGAGCAGCACGCTGAGCTATACCAAGAAATACAGCAATGATTCCT 545
DB      1317  GGCTCAAGTTAAAGTGGAGCAGCACGCTGAGCTGTACCAAGAAATACAGCAATTCCT 1376
QY      546  GGCCTACCTCAGCAACCGCTGCTGCCCCCAGTGACTACCGGAGTGGCTGCTCTTTG 605
DB      1377  GGCATACCTCAGCAACCGCTGCTGCCCCCAGTGACTACCGGAGTGGCTGATCTTTG 1436
QY      606  ATGTCACCGGAGTTGTGCGCGAGTGGCTGACCGCGCAGAGAGGCTATAGAGGGTTTCGCG 665
DB      1437  ATGTCACCGGAGTTGTGCGCGAGTGGCTGACCGCGCAGAGAGGAAATAGAGGGTTTCGCG 1496
QY      666  TCAGTGGCCACTCTCTCTGACAGCAAGCAATACACACTCCAGCTGGAATTAACGGGT 725
DB      1497  TTAGGGCCACTCTCTCTGACAGCAAGCAATACACACTCCAGCTGGAATTAACGGGT 1556
QY      726  TCAATTTGCGCGCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCTTCTCTGC 785
DB      1557  TCATACCGCGCGCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCTTCTCTGC 1616
QY      786  TCTCATGGCCACCGCGTGGAGAGGCGCCAGACCTGCAAGCTCCCGGACCGCGGAG 845
DB      1617  TTCTCATGGCCACCGCGTGGAGAGGCGCCAGACCTGCAAGCTCCCGGACCGCGGA - 1675
QY      846  CCCTGGATACCAACAGCTACCCATACAGCTGCCAGACTACGATCTCTGGCCCTGGATA 905
DB      1676  -----GCCCTGGACA 1685
QY      906  CCACTACTGCTTCACTCCACGAGAGAACTGCTGCTGCGGAGCTCTACATTTGACT 965
DB      1686  CCACTATTGCTTCACTCCACGAGAGAACTGCTGCTGCGGAGCTCTACATTTGACT 1745
```

```
QY      966  TCCGGAAGGACCTTGGCTGGAGTGGATTATGAAGCCCAAGGCTACCATGCAATTTCT 1025
DB      1746  TCCGGAAGGACCTTGGCTGGAGTGGATTATGAAGCCCAAGGCTACCATGCAATTTCT 1805
QY      1026  GCCTGGGGCCCTGCTCCCTACATCTGAGAGCTAGACACTAGTACACAGGTCCTGGGCTC 1085
DB      1806  GCCTGGGGCCCTGCTCCCTACATTTGGAGCTTGGAGCTGACGAGTACACAGGTCCTGGCC 1865
QY      1086  TGTAAACAGCAGCAACCCGGGCGCTGCGCGCGCGCTGCTGCTGCGCGAGGCGCTGG 1145
DB      1866  TGTAAACAGCAGCAACCCGGGCGCTGCGCGCGCGCTGCTGCTGCGCGAGGCGCTGG 1925
QY      1146  AGCCACTGCCATCTGCTACTAGTGGCGCGCAAGCCCAAGTGGAGCGAGCTGTCCAACA 1205
DB      1926  AGCCACTGCCATCTGCTACTAGTGGCGCGCAAGCCCAAGTGGAGCGAGCTGTCCAACA 1985
QY      1206  TGATCGTGGCTTCTGCAAGTGGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 1265
DB      1986  TGATCGTGGCTTCTGCAAGTGGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2045
QY      1266  GGCAGGCGCGCGCGCGCGCGCGCGCGCGCT-----CACCGGGGCTGTATTAAGGACA 1317
DB      2046  GGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCGCGCGCTGTATTAAGGACA 2105
QY      1318  TCGTGCCCAAGCCACTTGGGATCGATTAA 1349
DB      2106  CCGTGCCCAAGCCACTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2137

RESULT 9
OATCFBI
LOCUS      O.aries mRNA for transforming growth factor-beta 1.
DEFINITION X76916
ACCESSION X76916.1 GI:496648
VERSION   TGF-beta 1; transforming growth factor-beta 1.
KEYWORDS  Ovis aries (sheep)
SOURCE    Ovis aries
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovidae; Caprinae; Ovis.
REFERENCE 1 Woodall,C.J., McLaren,L.J. and Watt,N.J.
          Sequence and chromosomal localisation of the gene encoding ovine
          latent transforming growth factor-beta 1
          Gene 150 (2), 371-373 (1994)
JOURNAL   95121932
MEDLINE   7821809
REFERENCE 2 (bases 1 to 1173)
AUTHORS   Woodall,C.
TITLE     Direct Submission
JOURNAL   Submitted (24-DEC-1993) C. Woodall, Univ. of Edinburgh, Dept. of
          Veterinary Pathology, Sc. of Vet. Studies, Univ. of Edinburgh,
          Edinburgh EH9 1QH, UK
FEATURES  location/Qualifiers
           1..1173
           /organism="Ovis aries"
           /mol_type="mRNA"
           /db_xref="taxon:9940"
           1..1173
           /codon_start=1
           /product="transforming growth factor-beta 1"
           /protein_id="CAA54242.1"
           /db_xref="GI:496649"
           /db_xref="SWISS-PROT:P50414"
           /translation="MPPSGRLRLPLLLPLLLWLLMTPGRPVAGLSTKIDMELVKHK
           GIEAIRGQLSKRLASPPSGDVPGRPLPEAILALYNSTRDRVAESATPEPPAD
           IYAEVRLVAVEYGNKIYDKMSSSHISFNTSELREAVPEPVLISRADYRLRL
           KLVYORVELYQKYSNNRYSNRLAPSDSPWLSFDVTGVVRLTHREIEGR
           LSAHSCDSKONTLOVDINGFSRGRGLDATHGMNRPFLLMATPLEAQHLHSRHR
           RRLADTNYCFSSSTERNKCCVRLQYIDFRKDLGKWKIHERPKYHANFCLGCPYTWSLDT
           QYSKVILALYNHNPFGASAPCCVPQALEPLPIVYVGRKPKVQLSNMIVRSCKCS"
           mat_peptide 835..1170
```


QY	606	ATGTCACGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665			
DB	857	ATGTACCGGAGTTGTGCGGCAGTGGTTCAGCGCCGGAGGGAAATAGAGGGCTTTCGCC	916			
QY	666	TCAGTGGCCCACTCTTCTCTCTGACAGAAAGATAACACACTCCACGTCGGAATTAACGGGT	725			
DB	917	TTAGGCCCACTGCTCTCTGTGACAGAAAGATTACCACTGCAAGTGGACATCAACGGGT	976			
QY	726	TCAATTCTGGCCGCGGGGTGACCTGGCCACCACTTCACGGCATGAACCGGCCCTCTCTCTGC	785			
DB	977	TCACTACGGCGCGGAGTGAACCTGGCCACAATTCATGTCATGAACCGGCCCTTTCCTGC	1036			
QY	786	TCCTCATGGCCACCCCTCGAGAGGGCCACGACCTGCACAGCTCCCGGCACCGCCGAG	845			
DB	1037	TTCTCATGGCCACCCCACTGGAGAGGGCCCACTCTGCAAGCTCCCGGCACCGCCGA	1095			
QY	846	CCCTGGATACCAACAGACTACCCCATACGAGTGCACAGTACGCATCTCTGCCCTGGATA	905			
DB	1096	-----GCCCCTGGACA	1105			
QY	906	CAACTACTGTTTCAGCTCCACGGAGAGAACTGCTGGTGGCGGAGCTCTACATTGACT	965			
DB	1106	CAAATCTGTTTCAGCTCCACGGAGAGAACTGCTGGTGGCGAGCTGTATTGACT	1165			
QY	966	TCCGGAAGGACCTGGGCTGGAAGTGATTTCATGAACCCAAAGGCTTACCATGCCAATTCT	1025			
DB	1166	TCCGCAAGGACCTGGTGGAGTGGATCCACGAGCCCAAGGCTTACCATGCCAATTCT	1225			
QY	1026	GCCTGGGGCCCTGTCCTACATCTGAGGCTTAGCACTCAGTACAGCAAGTCTCTGGCTC	1085			
DB	1226	GCCTGGGGCCCTGTCCTACATCTGAGGCTGACACGCAAGTCTCTGGCTC	1285			
QY	1086	TGTACAACACGACAACCCGGGCGCGTGCSCGGCGCGTGTCTGCGTGCAGAGCGCTGG	1145			
DB	1286	TGTACAACACGACAACCCGGGCGCGTGCSCGGCGCGTGTCTGCGTGCAGAGCGCTGG	1345			
QY	1146	AGCCACTGCCCATCTGTACTAGTGGGCGCAAGCCCAAGTGGAGAGCTGTCCAACA	1205			
DB	1346	AGCCACTGCCCATCTGTACTAGTGGGCGCAAGCCCAAGTGGAGAGCTGTCCAACA	1405			
QY	1206	TGATGTGCTTCTCTGAAGTGCAGCTGAGGCCCGCCCGCCACACCGCCGCCACCC	1265			
DB	1406	TGATGTGCTTCTCTGAAGTGCAGCTGAGGCCCGCCCGCCACACCGCCGCCACCC	1465			
QY	1266	GGCAGGCCCGCCCGCCACCCCGCCCGCT-----CACCGGGCTGTATTAAAGGACA	1317			
DB	1466	GCCCGGCCCGCCCGCCACCCCGCCCGCTGTCTTGGCTTGGGGCTGTATTAAAGGACA	1525			
QY	1318	-TCGTGCCCAAGCCCACTGGGATCGATTAA	1349			
DB	1526	CCCGTGCCCCAAGCCCACTGGGGCCCAATAA	1558			
RESULT 13						
E03028						
LOCUS	E03028	1821 bp	RNA linear PAT 29-SEP-1997			
DEFINITION	DNA encoding human prepro TGF-beta1.					
ACCESSION	E03028					
VERSION	E03028.1 GI:2171250					
KEYWORDS	Jp 1991180192-A/1.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	Ohashi,H., Ishii,Y., Miyata,Y., Miyazono,K., Miyagawa,K. and Takaku,F.					
AUTHORS	1 (bases 1 to 1821)					
PRODUCTION OF HUMAN PRO-TGF-BETA1 BY GENETIC RECOMBINATION						
TITLE	Patent: JP 1991180192-A 1 06-AUG-1991;					
JOURNAL	KIRIN BREWERY CO LTD					
OS Homo sapiens (human)						
COMMENT	PN Jp 1991180192-A/1					

[illegible]

Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTAGTACTACCGGAGGCGGTGCTCTTG	605
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCGACAGTGTATCTTTTG	1106
Qy	606	ATGTACACGGAGTTGTCGGCAGTGGCTGACCGGAGAGAGCTATAGAGGTTTTTCGCC	665
Db	1107	ATGTACACGGAGTTGTCGGCAGTGGTTGAGCCGTGGAGGGGAATATGAGGGCTTTGCC	1166
Qy	666	TTAGTGCCCACTCTCTCTGACAGCAAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1167	TTAGTGCCCACTCTCTCTGACAGAGGAGATAACACACTGCAAGTGCACATCAACGGT	1226
Qy	726	TCAATTCTGGCCGCGGGGTGACTGSCCACCATTCAACGGATGAACCGCCCTTCCTGC	785
Db	1227	TCATACCGCGCGGAGGTGACTGCGCACCATTCATGGCAIGAACCGGCTTTCTCGC	1286
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGCCAGCAGCTGCACAGCTCCCGGACCGCGGAG	845
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGCCAGCAGCTGCAAGAGTCCCGGACCGCGGA	1345
Qy	846	CCCTGGATTACCAACAGCTACCCCATACGAGCTGCCAGCTACGCATCTCTGCCCCCTG	905
Db	1346	-----GCCCTGGACA	1355
Qy	906	CCAACTACTGTTTTCAGTCCACGGAGAGAAGTAAGTGTGCTGGCGCAGCTCTACATTG	965
Db	1356	CCAACTATTGTTTTCAGTCCACGGAGAGAAGTGTGCTGCTGGCGCAGCTGTACATTG	1415
Qy	966	TCGGAGGAGCTGGGCTGGAGTGGATTCATGAACCCGAAGGCTTACCATGCCAATTTCT	1025
Db	1416	TCGGAGGAGCTGGGCTGGAGTGGATTCACAGAGCCCAAGGCTTACCATGCCAATTTCT	1475
Qy	1026	GCTGGGGCCCTGCTCCCTACATCTGGAGCCTAGACACTCACTACAGCAAGTTCCTGGCTC	1085
Db	1476	GCTGGGGCCCTGCTCCCTACATCTGGAGCCTAGACACTCACTACAGCAAGTTCCTGGCTC	1535
Qy	1086	TGTACAAACGACACACCCGGGGCGCTGGCGGCGCGTGTGCTGGCGGAGCGGCTGG	1145
Db	1536	TGTACAAACGACATAACCCGGGGCGCTGGCGGCGCGTGTGCTGGCGGAGCGGCTGG	1595
Qy	1146	AGCCACTGCCATCGTGTACTACTGTCGGCGCGCAAGCCCAAGCTGGAGCAGCTGTCCAACA	1205
Db	1596	AGCCGCTGCCATCGTGTACTACTGTCGGCGCGCAAGCCCAAGCTGGAGCAGCTGTCCAACA	1655
Qy	1206	TGATGTCGGTTCCTGCAAGTGCAGCTGAGCGCCCGCGCGCCACAGCGCCCGCCACCC	1265
Db	1656	TGATGTCGGTTCCTGCAAGTGCAGCTGAGCGCCCGCGCGCCACAGCGCCCGCCACCC	1715
Qy	1266	GGCAGCGCGCGCCACCCCGCGCGCGCT-----CACCGGGCTGTATTAAAGACA	1317
Db	1716	GGCCGCGCGCCACCCCGCGCGCGCTGTGCTTGCCTTGGCCATGGGGCTGTATTAAAGACA	1775
Qy	1318	TCGTGCCCCCAAGCCC	1332
Db	1776	CGGTGCCCCCAAGCC	1790
RESULT	14		
LOCUS	BC022242	1746 bp	linear
DEFINITION	Homo sapiens, clone MGC:22008 IMAGE:4399762, mRNA, complete cds.		
ACCESSION	BC022242		
VERSION	BC022242.1		
KEYWORDS	GI:18490115		
SOURCE	MGC.		
ORGANISM	Homo sapiens (human)		
AUTHORS	Homo sapiens		
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1746)		
STRASBERG	Strausberg, R.		
DIRECT SUBMISSION	Submitted (01-FEB-2002) National Institutes of Health, Mammalian		
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer		

USA
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLML)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@axil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 27 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES	Location/Qualifiers
source	l..1746
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="MGC:22008 IMAGE:4399762"
	/tissue_type="Duodenum adenocarcinoma"
	/clone_lib="NIH_MGC_88"
	/lab_host="D410B"
	/note="Vector: pCMV-SPORT6"
CDS	370..1542
	/codon_start=1
	/product="Unknown (protein for MGC:22008)"
	/protein_id="AA42242.1"
	/db_xref="GI:18490116"
	/translation="MPSPGURILLLLPLWLVLTPGRPAAGLSTCKTIDMBLVRRKRIEAIHQGLSKRLASPPSQGVPPGLPEVALIYNSTRDVAGESASPEPEPADKLAVKEIVRLVYTHNEIYDKFKQSHSIYMFNTSELREAVPEPVLSSAEELRLKLVAKEOHVELVYQXSNNSRYSLSNRLAPSDSEWISFDVTGVVRWLSRGSGIEGFRLSAHCSDSDNTLOYDIHTFTGGRLDLATIHGMNRPFLLMATPLERAQHQLQSSRHRAALDNYCPSFSSTKNCVCVPIAYIDFKDLQGWKIHEPKGYHANFCILGPCPYTWLDTOYSKVIALYNQHPGASAAPCCVPQALEPLPIYYVGKKPKVEQLSNHIVRSKCS"

BASE COUNT	376 a	612 c	472 g	286 t
ORIGIN				
Query Match	71.1%; Score 967.8; DB 9; Length 1746;			
Best Local Similarity	84.7%; Pred. No. 1.9e+165;			
Matches 1146; Conservative); Mismatches 147; Indels 60; Gaps 3;			
Qy	6	CGAGATGGCGCTTCGGGCTCTTGCGGTCTTGCGCTGCTGCCGCTGCTGTGGCTGC	65	
Dd	365	CCCCCATGCGCGCCCTCGGGGCTCGGGGTGCTGCTGCTGCTACCCTGCTGTGGCTAC	424	
Qy	66	TAGTGTGAGCGCTTGGCGGCGCGCGCCCGGACTGTCCACTGCAAGCACATCGACATGG	125	
Dd	425	TGGTGTGACGCTGGCGCGCGCGCGGGACTATCCACTGCAAGCATATCGACATGG	484	
Qy	126	AGCTGTGAAGCGAAGCGATATGAGGCGCATTCGCGGCGAGATTCTGTCCAAGTTCCGGC	185	
Dd	485	AGCTGTGAAGCGAAGCGATATGAGGCGCATTCGCGGCGAGATTCTGTCCAAGTTCCGGC	544	
Qy	186	TTGCCAGCCCCCAGGCAAGGAGACGTGCGCGCGCGCGCGCTGCCTGAGGCGACTATGG	245	
Dd	545	TCGCCAGCCCCCAGGCAAGGAGACGTGCGCGCGCGCGCGCTGCCTGAGGCGCGCTGCTG	604	
Qy	246	CTCTTTACAACAGTACCGCGACCGCGGTAGCGGGGGAAGGTGTGCAACCGGAGCCGAGC	305	
Dd	605	CCCTGTACACAGCACCGCGACCGGTGCGCGGGGAGAGTGCAGAACCGGAGCCGAGC	664	
Qy	306	CAGAGCGGACTACTACGCCAAGGAGGTACCCCGGTGCTTAATGGTGGAAAGCGGCAACC	365	
Dd	665	CTGAGCGGACTACTACGCCAAGGAGGTACCCCGGTGCTTAATGGTGGAAAGCGGCAACC	724	

Search completed: October 9, 2003, 17:40:16
Job time : 5264 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2003, 14:56:03 ; Search time 429 Seconds
(without alignments)

8563.951 Million cell updates/sec

Title: US-10-017-372E-36

Perfect score: 1361

Sequence: 1 tggtaacagatgagccctt.....cattaaagcgcgcgact 1361

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03:.*
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1211.8	89.0	1326	24	AA022696 Porcine transforming growth factor beta 1 (TGF-beta1) cDNA.
2	979.6	72.0	25	ABQ76674	Porcine transforming growth factor beta 1; TGF-beta1; gene therapy; IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive; multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus; diabetes mellitus; sarcoidosis; psoriasis; dermatological; ss.
3	978.8	71.9	2537	7	PAN60972 Sus scrofa.
4	978.8	71.9	2537	11	AA003301 Location/Qualifiers
5	978.8	71.9	2537	11	AA003301 16..1188
6	978.8	71.9	2537	17	AA002814 /product- "Porcine TGF-beta1 mutant protein"
7	977.2	71.8	2537	15	AAQ56923 WO200181404-A2.
8	974	71.6	2537	19	AAV52933 01-NOV-2001.

9	971	71.3	1561	11	AAQ03268	Simian transformin
10	969.4	71.2	2742	22	AA158342	Human polynucleoti
11	968.6	71.2	1559	13	AAQ20389	Sequence encoding
12	968.2	71.1	1821	12	AAQ13392	Human pro-TGF-beta
13	967.6	71.1	1560	9	AAAN81084	Coding sequence of
14	967.6	71.1	1560	11	AAQ03508	Simian Transformin
15	965.8	71.0	2745	16	AA105876	cDNA encoding tran
16	965.8	71.0	2745	22	AAH28216	Nucleotide sequenc
17	957.8	70.4	1303	11	AAO09317	Monkey transformin
18	956.8	70.3	4105	15	AAQ55624	TGFbeta1 5'-UTR-CD
19	954.8	70.2	1571	11	AAQ03269	Human transforming
20	949.8	69.8	1569	9	AAAN81085	Coding sequence of
21	946.6	69.6	1561	11	AAQ03509	Human Transforming
22	940.6	69.1	1561	13	AAQ04908	Sequence encoding
23	937.4	68.9	1561	13	AAQ29177	TGF-beta 1/beta 2
24	918.2	67.5	1176	25	ABV75391	TGFbeta1 Arg25Pro pol
25	916.6	67.3	1176	24	AB235738	Human TGF beta 1 p
26	916.6	67.3	1176	24	ABX09981	Human TGFbeta1 DNA
27	916.6	67.3	1176	24	ABX09981	Human TGF beta 1 D
28	916.6	67.3	1176	24	ABX09981	Human polynucleoti
29	916.6	67.3	1176	25	ABV75392	TGFbeta1 Arg25Pro pol
30	896	65.8	1565	13	AAQ29178	TGF-beta 1. Homo
31	797.4	58.6	2765	22	AA160128	Human polynucleoti
32	776.2	57.0	2208	13	AAQ20291	Sequence encoding
33	774.6	56.9	2208	11	AAQ03510	Human Transforming
34	774.6	56.9	2208	11	AAQ03511	Hybrid transformin
35	773	56.8	2207	11	AAQ05127	Human TGF-Beta1/TG
36	761.6	56.0	2217	10	AAAN90768	Sequence of human
37	755.4	55.5	2200	16	AA104115	Simian-human hybri
38	684.6	50.3	2773	23	AA584421	DNA encoding novel
39	678.8	49.9	834	12	AAQ12192	Sequence encoding
40	654	48.1	1376	24	ABK90341	DNA encoding LAP-m
41	652.4	47.9	1389	24	ABK90344	DNA encoding LAP-h
42	588.2	43.2	1352	24	ABK90342	DNA encoding MIFNB
43	586.6	43.1	1350	24	ABK90343	DNA encoding huIFN
44	522.6	38.4	875	23	AA570979	DNA encoding novel
45	357.6	26.3	650	24	ABK84023	Human cDNA differe

ALIGNMENTS

RESULT 1	
AA022696	
ID	AA022696 standard; cDNA; 1326 BP.
XX	
AC	AA022696;
XX	
DI	26-FEB-2002 (first entry)
XX	
DE	Porcine transforming growth factor beta 1 (TGF-beta1) cDNA.
XX	
KW	Porcine; transforming growth factor beta 1; TGF-beta1; gene therapy;
KW	IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;
KW	multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;
KW	diabetes mellitus; sarcoidosis; psoriasis; dermatological; ss.
OS	Sus scrofa.
XX	
FH	Key
CDS	Location/Qualifiers
FT	16..1188
ET	/*tag-
FT	/product- "Porcine TGF-beta1 mutant protein"
XX	
PN	WO200181404-A2.
XX	
PD	01-NOV-2001.
XX	
PF	20-APR-2001; 2001WO-US12980.
XX	
PR	20-APR-2000; 2000US-199014P.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Strober W, Nakamura K, Kitani A, Fuss IJ;
XX WPI; 2002-026155/03.
DR P-PSDB; AAE13596.

XX Composition for treating autoimmune diseases e.g. inflammatory bowel
PT disease in humans, comprises vector containing transforming growth
PT factor-beta under the control of inducible promoter -

XX Claim 1; Fig 1; 78pp; English.

XX The invention relates to a composition containing a vector comprising a
CC gene encoding a regulatory transcription factor under the control of a
CC promoter encoding a transforming growth factor-beta (TGF-beta). The
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2
CC or TGF-beta3, its variants or homologues, by transfecting a cell which
CC is part of a host suspected of having an autoimmune disease, especially
CC inflammatory bowel disease (IBD), under conditions such that the
CC polypeptide encoded by the nucleic acid sequence in the vector is
CC expressed. The vector is delivered using a delivery system. The delivery
CC of the vector results in substantial elimination of symptoms of the
CC autoimmune disease and increased production of IL-10 by the host. The
CC composition is useful for treating various diseases with an autoimmune
CC component such as multiple sclerosis, rheumatoid arthritis, systemic
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis
CC and psoriasis, and also for assaying the expression of a gene in a cell.
CC The vector is further useful for screening of the effect of test
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.
CC The present sequence is a cDNA encoding porcine TGF-beta1 mutant.

XX Sequence 1326 BP; 263 A; 438 C; 392 G; 233 T; 0 other;

Query Match 89.0%; Score 1211.8; DB 24; Length 1326;
Best Local Similarity 95.6%; Pred. No. 3.2e-235;
Matches 1299; Conservative 0; Mismatches 7; Indels 53; Gaps 3;

QY 1 TGGTACCGAGATGGCGCTTGGGCTGCGGCTCTTGGCGTGTGCTGCGCTGCTGCTG 60
DB 6 TGGTACCGAGATGGCGCTTGGGCTGCGGCTCTTGGCGTGTGCTGCGCTGCTGCTG 65
QY 61 GCTGCTAGTGTGAGCGCTGCGCGCGCGCGCGGAGTGTCCACCTGCGAAGACCATCGA 120
DB 66 GCTGCTAGTGTGAGCGCTGCGCGCGCGCGCGGAGTGTCCACCTGCGAAGACCATCGA 125
QY 121 CATGAGCTGTGTGAAGCGGAGCGCATCGAGGCCATTGCGCGCGGAGATTCGTCCAAAGCT 180
DB 126 CATGAGCTGTGTGAAGCGGAGCGCATCGAGGCCATTGCGCGCGGAGATTCGTCCAAAGCT 185
QY 181 TCGGCTTGGCAGCGCGCGCGCGCGCGCGCGGAGTGTCCACCTGCGAAGACCATCGA 240
DB 186 TCGGCTTGGCAGCGCGCGCGCGCGCGCGGAGTGTCCACCTGCGAAGACCATCGA 245
QY 241 ACTGCTCTTTACACAGTACCGCGACCGGCTAGCGCGGGAAGTGTGCAACCGGAGCC 300
DB 246 ACTGCTCTTTACACAGTACCGCGACCGGCTAGCGCGGGAAGTGTGCAACCGGAGCC 305
QY 301 CGAGCCAGAGCGGCGACTACTAGCCCAAGAGGTACCGCGGTGTATATGCTGTTCAACAC 360
DB 306 CGAGCCAGAGCGGCGACTACTAGCCCAAGAGGTACCGCGGTGTATATGCTGTTCAACAC 365
QY 361 CAACCAATCTATGATAATTAACAGGCGACCCCGACAGCTTATATGCTGTTCAACAC 420
DB 366 CAACCAATCTATGATAATTAACAGGCGACCCCGACAGCTTATATGCTGTTCAACAC 425
QY 421 GTCGGAGCTCCGGAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGGCT 480
DB 426 GTCGGAGCTCCGGAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGGCT 485
QY 481 GCTGAGGCTCAAGTTAAAGTGGAGCGAGCGTGTGAGCTATACCAAGAAATACAGCAATGA 540
DB 486 GCTGAGGCTCAAGTTAAAGTGGAGCGAGCGTGTGAGCTATACCAAGAAATACAGCAATGA 545

QY 541 TTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCGCCAGTGACTCACCGGAGTGGCTGTC 600
DB 546 TTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCGCCAGTGACTCACCGGAGTGGCTGTC 605
QY 601 CTTTGATGTCACCGGAGTGTGGCGGAGTGGCTGACCCCGCAGAGGCTATAGAGGTTT 660
DB 606 CTTTGATGTCACCGGAGTGTGGCGGAGTGGCTGACCCCGCAGAGGCTATAGAGGTTT 665
QY 661 TCGGCTCAGTCCCGACCTCTCTCTGACAGCAAGAGATAACACACTCCACCTGCAATTA 720
DB 666 TCGGCTCAGTCCCGACCTCTCTCTGACAGCAAGAGATAACACACTCCCGTGGAAATTA 725
QY 721 CGGGTTCAATTTCGGCCCGCGGGGTGACCTTGGCCACCATTCACGGCATGAACCGGCTT 780
DB 726 CGGGTTCAATTTCGGCCCGCGGGGTGACCTTGGCCACCATTCACGGCATGAACCGGCTT 785
QY 781 CTTGCTCTCTATGGCCACCGCGCTGGAGAGGCGCCACAGCTGCACAGCTCCCGGACCG 840
DB 786 CTTGCTCTCTATGGCCACCGCGCTGGAGAGGCGCCACAGCTGCACAGCTCCCGGACCG 845
QY 841 CGAGCCCTTGGATACCAACAGCTACCCATACGAGCTGCCAGACTACGCAATCTCTG 900
DB 846 CGGA-----GCGCT 854
QY 901 GGATACCAACTACTGCTTCACTCCAGGAGCAAGACTGCTGCGCGCAGCTCTACAT 960
DB 855 GGATACCAACTACTGCTTCACTCCAGGAGCAAGACTGCTGCGCGCAGCTCTACAT 914
QY 961 TGACTTCCGGAAGGACCTGGGCTGGAGTGGATTTCATGAACCCAAAGGCTTACCATGCCAA 1020
DB 915 TGACTTCCGGAAGGACCTGGGCTGGAGTGGATTTCATGAACCCAAAGGCTTACCATGCCAA 974
QY 1021 TTTCTGCTTGGGCGCTCTGCTTACATCTGAGGCTACACACTCAGTACAGCAAGGTCCT 1080
DB 975 TTTCTGCTTGGGCGCTCTGCTTACATCTGAGGCTACACACTCAGTACAGCAAGGTCCT 1034
QY 1081 GGCTCTGTACAAACAGCAACCGCGCGCTGCGGCGCGCTGCTGCTGCGCGCAGGC 1140
DB 1035 GGCTCTGTACAAACAGCAACCGCGCGCGCTGCGGCGCGCTGCTGCTGCGCGCAGGC 1094
QY 1141 GCTGAGGCACTGCCCATCTGCTACTAGTGGCGCGCAAGCCCAAGGTGGAGCAGTCTC 1200
DB 1095 GCTGAGGCACTGCCCATCTGCTACTAGTGGCGCGCAAGCCCAAGGTGGAGCAGTCTC 1154
QY 1201 CAACATGATGCTGCTTCTGCAAGTGCAGCTGAGGCGCGCGCGCGCGCGCGCGCGCC 1260
DB 1155 CAACATGATGCTGCTTCTGCAAGTGCAGCTGAGGCGCGCGCGCGCGCGCGCGCGCC 1213
QY 1261 CACCGCGGAGCC 1320
DB 1214 CACCGCGGAGCC 1273
QY 1321 TGCCCAAGCGCGCGCTGGGATCGATTAAAGCGCGCGCA 1359
DB 1274 TGCCCAAGCGCGCGCTGGGATCGATTAAAGGTGGAGAGA 1311

RESULT 2
ABQ76674
ID ABQ76674 standard; DNA: 2527 BP.

XX AC ABQ76674;

XX XX 26-MAR-2003 (first entry)

XX Androgen receptor signalling pathway-associated DNA E00973.

DE DE Androgen receptor; transactivation; modulator; Smad4; Akt; TGF-B;
KW signal transduction pathway; transforming growth factor-B; phosphatase;
KW tensin; cytosolic; antiproliferative; cellular proliferation; cancer;
KW E00973; ds.
XX OS Synthetic.

XX AC AAQ03301;
 XX AC 25-MAR-2003 (updated)
 XX DT 05-AUG-1990 (first entry)
 XX DE cDNA encoding human pre-transforming growth factor-beta-1 (pre-TGF-beta-1).
 XX DE 1).
 XX DE
 XX KW Transforming growth factor-beta-1 (TGF-beta-1);
 XX KW neoplastic cell line inhibition;
 XX KW EGF-potentiased anchorage-independent growth;
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 842..2014
 XX FT /tag- a
 XX FT mat_peptide 1676..2011
 XX FT /tag- b
 XX FT misc_difference 37..113
 XX FT /tag- c
 XX FT misc_feature 2015..2100
 XX FT /tag- d
 XX FT /note="G-C rich sequence
 XX FT and a downstream TATA-like sequence"
 XX PN US4886747-A.
 XX XX 12-DEC-1989.
 XX XX
 XX PF 13-MAR-1987; 87US-0025423.
 XX PR 13-MAR-1987; 87US-0025423.
 XX XX (GETH) GENENTECH INC.
 XX PA Derynck RMA, Goeddel DV;
 XX PI WPI; 1990-051338/07.
 XX DR P-ESDB; AAR05258.
 XX PT Nucleic acid encoding transforming growth factor-beta -
 XX PT cloned into expression vectors for expression in eukaryotic host
 XX PT cells for therapeutic use
 XX PS Disclosure; Fig 1b; 28pp; English.
 XX CC It was obtained by an analysis of several overlapping cDNAs and gene
 XX CC fragments, leading to the detn. of a continuous sequence corresp. to the
 XX CC TGF-beta-1 precursor mRNA. It is useful in constructing vectors that
 XX CC encode biologically active transforming growth factor (TGF-beta),
 XX CC operably linked to DNA that encodes a secretory leader (SL). It, or a
 XX CC nucleic acid capable of hybridising with it, can also be labelled and
 XX CC used in diagnostic assays for DNA or mRNA encoding TGF-beta or related
 XX CC proteins.
 XX CC (Updated on 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;
 Query Match 71.9%; Score 978.8; DB 11; Length 2537;
 Best Local Similarity 84.8%; Pred. No. 3.4e-188;
 Matches 1146; Conservative 0; Mismatches 147; Indels 59; Gaps 2;
 OY 6 CCGAGATCGCCCTCGGGCTGCGGCTCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTG 65
 DB 837 CCCCCATCCCGCTCGGGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCT 896
 OY 66 TAGTGCTGACGCGCTCGGGCGCGCGGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTG 125
 DB 897 TGGTGCTGACGCGCTCGGGCGCGCGGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTG 956

OY 126 AGCTGCTGAAGCGGAAGCGGCATCCAGGCCATTCCGGCCAGATTCTGTCCAAAGTTCCGGC 185
 DB 957 AGCTGCTGAAGCGGAAGCGGCATCCAGGCCATCCCGGCCAGATCTGTCCAAAGTTCCGGC 1016
 OY 186 TTGCCAGCCCCCGAGCCAGGGGACGTGCGCGCCGCGGCTGCTGCTGAGGACGCTACTGG 245
 DB 1017 TCGCAGACCCCCCGAGCCAGGGGAGGTGCGCGCCGCGGCTGCTGCTGAGGCGCTGCTG 1076
 OY 246 CTCCTTACACAGTACCCCGCAGCGGGTAGCCGGGGAAGTGTCTGAAACCGGAGCCGAGC 305
 DB 1077 CCCTGTACAACAGCACCCCGCAGCGGGTGGCGGGGAGGTGCAGAACCGGAGCCGAGC 1136
 OY 306 CAGAGCGGAGCTACTACGCCAAGGAGGTACCCCGGCTGCTAATGGTGAAGAGCGGCAACC 365
 DB 1137 CTGAGCGCGGCTACTACGCCAAGGAGGTACCCCGGCTGCTAATGGTGAAGAGCGGCAACC 1196
 OY 366 AATCTATGATAAATTCAGGGGCACCCCGCCACAGCTTATATATCTGTTCAACACGTCGG 425
 DB 1197 AATCTATGACAAAGTTCAGGAGAGTACACACAGCATATATATCTTCAACACATCAG 1256
 OY 426 AGCTCCGGGAAGCGGTGCCGGAACCTGTTATGCTCTCTCGGGCAGAGCTGCGCTGTGTA 485
 DB 1257 AGCTCCGGAAGCGGTACTGAAACCGGTGTGCTCTCCCGGCGAGAGTGGCTGCTGTA 1316
 OY 486 GCCTCAAGTFAAAGTGGAGCAGCGTGGAGCTATACCAAAATACAGCAATGATTCCT 545
 DB 1317 GGCTCAAGTFAAAGTGGAGCAGCGTGGAGCTGATCCAGAAATACAGCAATGATTCCT 1376
 OY 546 GGCGTACTCTAGCAACCGGCTGCTGGCCCGCCAGTACTCACCGGAGTGGTGTCTTTG 605
 DB 1377 GCGGATACCTCAGCAACCGGCTGCTGGCCACCGAGCTCGCCAGAGTGGTATCTTTG 1436
 OY 506 ATGTCACCGGAGTTGTGCGGAGTGGCTGACCCCGCAGAGAGCTATAGAGGTTTTCGCC 665
 DB 1437 ATGTCACCGGAGTTGTGCGGAGTGGCTGACCCCGCAGAGAGTATAGAGGTTTTCGCC 1496
 OY 566 TCAGTGGCCCACTCTTCTCTGACAGCAAGATTAACACACTCCACCGTGAATTAACGGGT 725
 DB 1497 TTAGCGCCCACTCTTCTCTGACAGCAAGATTAACACACTCCACCGTGAATTAACGGGT 1556
 OY 726 TCAATCTGCGCGCGGCTGACCTGGCCACCATTCACGGCATGAACCGGCTTCTCTGC 785
 DB 1557 TCACCTACCGCGCGCGGCTGACCTGGCCACCATTCATGGCATGAACCGGCTTCTCTGC 1616
 OY 786 TCTCATGCGCACCCCGCTGGAGAGGCGCCAGCAGCTGCACAGCTCCCGGAGCGCGGAG 845
 DB 1617 TTCTCATGCGCACCCCGCTGGAGAGGCGCCAGCAGCTGCAGAGCTCCCGGAGCGCGGA 1675
 OY 846 CCCTGGATACCAACAGCTACCCATACGAGTCCAGACTACGCATCTCTGGCCCTGGATA 905
 DB 1676 -----GCGCTGGACA 1685
 OY 906 CCAACTACTGCTTACGCTCCACGAGAGAGTCTGCTGGTGGCGAGCTCTACATTGACT 965
 DB 1686 CCAACTATTGCTTACGCTCCACGAGAGAGTCTGCTGGTGGCGAGCTGTACATTGACT 1745
 OY 966 TCGGAAGGACCTGGGCTGGAAGTGGATTGATTAACCAAGGCTTACCATGCAATTTCT 1025
 DB 1746 TCGGCAAGGACCTGGGCTGGAAGTGGATTGATTAACCAAGGCTTACCATGCAATTTCT 1805
 OY 1026 GCGTGGGCGCTGTCCTACATCTGGAGCTTACAGCTAGACACTAGTACAGCAAGTCTCGGTC 1085
 DB 1806 GCGTGGGCGCTGTCCTACATCTGGAGCTTACAGCTAGACACTAGTACAGCAAGTCTCGGTC 1865
 OY 1086 TGTACAAACAGCAGCAACCGGCGCTGCGCGGCGCTGCTGCTGCGGAGGCGCTGG 1145
 DB 1866 TGTACAAACAGCAGCAACCGGCGCTGCGCGGCGCTGCTGCTGCGGAGGCGCTGG 1925
 OY 1146 AGCCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
 DB 1926 AGCCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1985
 OY 1206 TGAATCTGCTTCTCTGCAAGTGCAGTGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1265


```

Db 1676 -----GCCCTGGACA 1685
QY 906 CCAACTACTGCTTCACTCCACGGAGAGAAGAACTGCTGCTGGCAGCTCTACATTTGACT 965
Db 1686 CCAACTATTGCTTCACTCCACGGAGAGAAGAACTGCTGCTGGCAGCTGTACATTTGACT 1745
QY 966 TCCGGAAGGACTGGCTGCAAGTGGATTTCATGAACCAAGGCTTACCATCCCAATTTCT 1025
Db 1746 TCCGGAAGGACTGGCTGCAAGTGGATTTCATGAACCAAGGCTTACCATCCCAATTTCT 1805
QY 1026 GCTTGGGGCCCTGTCCTTACATCTTGAGCCCTAGACACTCAGTACAGCAAGTCTCTGGCTC 1085
Db 1806 GCTTGGGGCCCTGTCCTTACATCTTGAGCCCTAGACACTCAGTACAGCAAGTCTCTGGCTC 1865
QY 1086 TGTAAACCAACGACAAACCCGGCGCTGGCGGCGCCGTGCTGGTGGCCGAGGGCTGG 1145
Db 1866 TGTAAACCAACGACAAACCCGGCGCTGGCGGCGCCGTGCTGGTGGCCGAGGGCTGG 1925
QY 1146 AGCCACTGCCATCTGTACTACGTGGGCCCAAGCCCAAGGTGGAGCAGCTGTCCAAACA 1205
Db 1926 AGCCGCTGCCATCTGTACTACGTGGGCCCAAGCCCAAGGTGGAGCAGCTGTCCAAACA 1985
QY 1206 TGATCGTGGTTCCTGCAAGTGCAGCTGAGGCCCCCGCCCGCCACAGCCCGCCCAACC 1265
Db 1986 TGATCGTGGTTCCTGCAAGTGCAGCTGAGGCCCCCGCCCGCCACAGCCCGCCCAACC 2045
QY 1266 GGCAGCGCGCGCCACCCCGCCCGCT-----CACCGGGCTGTATTAAAGGACA 1317
Db 2046 GCGCGGCGCGCCACCCCGCCCGCTGCTTGCCTATGGGCGCTATTAAAGGACA 2105
QY 1318 TGGTGGCCCAAGCCCACTTGGGATCCATTA 1349
Db 2106 CCGTSCCCCAAGCCCACTTGGGCGCCCATTA 2137

RESULT 6
AAT15720
ID AAT15720 standard; cDNA; 2537 BP.
AC AAT15720;
XX
XX 25-MAR-2003 (updated)
DT 24-JUL-1997 (revised)
DT 25-JAN-1980 (first entry)
XX
DE Pre-transforming growth factor beta 1 cDNA.
XX
XX transforming growth factor beta 1; wound healing;
KW recombinant production; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..841
FT 5'UTR
FT misc_feature
FT 37..113
FT /*tag= a
FT /*tag= b
FT /note= "GC-rich region forms stable hairpin loops;
FT similar to structural organisation of c-myc RNA,
FT could play role in mRNA stability or in
FT regulation of transcription"
FT CDS 842..2014
FT /*tag= c
FT /product= pre-TGF_beta_1
FT mat_peptide 1676..2011
FT /*tag= d
FT /product= mature_TGF_beta_1
FT repeat_region 2015..2100
FT /*tag= e
FT /note= "GC-rich region; possibly responsible for the
FT fact 3'UTR of mRNA could not be cloned as cDNA;
FT may be important for transcription efficiency"
FT repeat_unit 2019..2023

```

```

FT TATA_signal /*tag= f
FT 2094...2100
FT /*tag= g
FT /note= "TATA-like sequence; no evidence that this
FT functions a promoter"
FT polyA_signal 2514..2520
FT /*tag= h
FT misc_signal 2529..2536
FT /*tag= i
FT /note= "consensus sequence immediately precedes
FT polyA-tail (Benoist et al)"
FT US5482851-A.
XX
XX 09-JAN-1996.
XX
XX 05-NOV-1993; 93US-0147364.
XX
XX 13-MAR-1987; 87US-0025423.
XX 22-MAR-1985; 85US-0715142.
XX 04-AUG-1989; 89US-0389929.
XX 04-MAR-1992; 92US-0845893.
XX 05-NOV-1993; 93US-0147364.
XX (GETH ) GENENTECH INC.
XX
XX Derynck RNA, Goeddel DV;
XX WP; 1996-076891/08.
XX P-PSDB; AAR90827.
XX
XX New recombinant human transforming growth factor-beta prods. - produced
XX using Chinese hamster ovary cells, for use in diagnostic applications
XX or in therapy
XX
XX Example 3; Fig 1; 26pp; English.
XX
XX The cDNA encodes the pre-transforming growth factor (TGF) beta 1 protein.
XX The nucleotide sequence was obtd. by an analysis of several overlapping
XX cDNAs and gene fragments. The DNA is useful for the recombinant
XX production of TGF beta 1, which can be used in, e.g. wound healing.
XX (Revised entry submitted to correct sequence analysis breakdown.)
XX (Updated on 25-MAR-2003 to correct pf field.)
XX
SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 71.9%; Score 978.8; DB 17; Length 2537;
Best Local Similarity 84.8%; Pred. No. 3.4e-188;
Matches 1146; Conservative 0; Mismatches 147; Indels 59; Gaps 2;

QY 6 CCGAGATGGCGCTTCGGGCTCGGCTCTTGCGGCTGCTGCGCGCTGCTGCGCTG 65
Db 837 CCCCCATGCCGCCCTCCGGGCTGGGCTGCTGCGCTGCTGCTGCTGCTGCTGCTG 896
QY 66 TAGTGTGACGCTTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
Db 897 TGGTGTGACGCTTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 956
QY 126 AGCTGGTGAACGCGAAGCGCATCCAGGCCATTCAGGCCAGATTCCTCCAACTTCGGC 185
Db 957 AGCTGGTGAACGCGAAGCGCATCCAGGCCATTCAGGCCAGATTCCTCCAACTTCGGC 1016
QY 186 TTGCAGCCCCCGAGCGAGGGGACGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 245
Db 1017 TCGCCAGCCCCCGAGCGAGGGGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1076
QY 246 CTCTTTAAACAGTACCCCGGACCGGGTACCCGGGAAAGTGTGCAACCCGAGCCCGAGC 305
Db 1077 CCCTGTACACAGCACCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1136
QY 306 CAGAGCGGGACTACTACGCCAAGGAGGTACCCCGCGTGTCTAATGGTGAAGGCGCAACC 365
Db 1137 CTGAGCGCGACTACTACGCCAAGGAGGTACCCCGCGTGTCTAATGGTGAAGGCGCAACC 1196

```


957 AGCTGGTGAAGCGAAGCGCATCGAGGCAATCCGCGCCAGATCCTGCAAGCTGGGC 1016
 186 TTGCGAGCCCGCCAGCAGCGAGGGGACGTGCGCGCCGCGCGCTGCTGAGGCACTACTGG 245
 1017 TCGCGAGCCCGCCAGCAGCGAGGGGAGGTGCGCGCCGCGCGCTGCGCGAGCGGTGCTCG 1076
 246 CTCCTTTACACAGTACCGCGACCGGGTAGCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
 1077 CCCTGTACAACAGCACCGCGACCGGGTGGCGGGGAGAGTGCAGAACGGGAGAGCCCGAGC 1136
 306 CAGAGCGGAGCTACTACGCCAAGGAGGTCAACCGCGTGTAAATGGTGGAAAGGGCAAC 365
 1137 CTGAGCGGAGCTACTACGCCAAGGAGGTCAACCGCGTGTAAATGGTGGAAACCCACAACG 1196
 366 AAATCTATGATAAATTCAGGGCACCGCCACAGCTTATATATGCTGTTCACACGTCGG 425
 1197 AAATCTATGACAAGTTCAAGTAGAGTACACACAGCATATATATGCTTCAACACATCAG 1256
 426 AGCTCCGGGAAGCGGTGCGGTAACCTGTATTGCTCTCGGGCAGAGCTGGCGCTGCTGA 485
 1257 AGCTCCGAGAAGCGGTACCTTAACCGGTGTTGCTCCCGGCGAGCTGCGTCTGCTGA 1316
 486 GCCTCAAGTTAAAGTGGAGTAGCACGCTGGAGCTATACCAAAATACAGCAATGATTCT 545
 1317 GGCTCAAGTTAAAGTGGAGTAGCACGCTGGAGCTATACCAAAATACAGCAATGATTCT 1376
 546 GCGGTACTCTCAGCAACCGGTGCTGGCGCCCGAGTACTACCGGAGTGGCTGTCTTTG 605
 1377 GCGGTACTCTCAGCAACCGGTGCTGGCACCCAGCACTCGCCAGAGTGGTGTATCTTTG 1436
 606 ATGTACACGGAGTTGTCGGGAGTGGCTGACCGCGAGAGGCTATAGAGGTTTTCGCC 665
 1437 ATGTACACGGAGTTGTCGGGAGTGGCTGACCGCGAGAGGCTATAGAGGTTTTCGCC 1496
 666 TCAATCTGGCCACTTTCCTCTTACAGCAAGATAACACACTCCAGCTGGAAATTAACGGGT 725
 1497 TTAGCGCCACTGCTCTGTACAGCAGGATACACACTCAAGTGCATCAACGGGT 1556
 726 TCAATCTGGCCGGGTACCTGGCGCCAGCTATCAAGGCTGATCAAGCGGCTTCCTGC 785
 1557 TCATACCGCGCGCGAGGTACCTGGCGCCAGCTATCAAGGCTGATCAAGCGGCTTCCTGC 1616
 786 TCCTCATGCGCACCCCGTGTAGAGGGCCAGCAGCTGCACAGCTCCCGGCGACCGCGAG 845
 1617 TTCATGCGCACCCCGTGTAGAGGGCCAGCAGCTGCACAGCTCCCGGCGACCGCGA- 1675
 846 CCCTGGATACCAACAGCTACCGATACGAGCTGCCAGACTACGATCTCTGCGCCCTGGATA 905
 1676 -----CCCTGGACA 1685
 906 CCAACTACTGCTTACGCTCCACGGAGAAGAACTGCTGGTGGCGAGCTCTACATTGACT 965
 1686 CCAACTATTGCTTACGCTCCACGGAGAAGAACTGCTGGTGGCGAGCTGTACATTGACT 1745
 966 TCCGAAGACCTCGGCTGGAGTGGATTATGAACCCAAAGGCTACCATGCCAATTTCT 1025
 1746 TCCGAAGACCTCGGCTGGAGTGGATTATGAACCCAAAGGCTACCATGCCAATTTCT 1805
 1026 GCTTGGGCGCTGCTCCCTACATCTCGAGCTAGACACTCAGTACAGCAAGTCTGCGCTC 1085
 1806 GCTTGGGCGCTGCTCCCTACATTTGGAGCTTGGACCGAGTACAGCAAGTCTGCGGCC 1865
 1086 TGTACAACAGCACACCGGCGCGGTGGCGGCGCTGCTGGTGGCGGAGGCGCTGG 1145
 1866 TGTACAACAGCATACCCGCGCGCTCGCGGCGCGCTGCTGGTGGCGGAGGCGCTGG 1925
 1146 AGCCACTGCCATCTGTACTACGTGGCGCGCAAGCCAAAGTGAGAGCTGTGCCAACA 1205
 1926 AGCCGCTGCCATCTGTACTACGTGGCGCGCAAGCCAAAGTGAGAGCTGTGCCAACA 1985
 1206 TCATCGTGGTTCCTGCAAGTGCAGCTGAGCGCCCGCCCGCCAGCGCGCCACCC 1265

1986 TGATCTGCTGCTCTCTGCAAGTGCAGTGCAGTCCCGCCCGCCCGCCCGCCCGCGCAG 2045
 1266 GGCAGGCCCGGCCCGCCACCCCGCCCGCT-----CACCGGGGCTGTATTAAAGGACA 1317
 2046 GCGCGGCCCGCCACCCCGCCCGCGCTGCTTGCATGGGGCTGTATTAAAGGACA 2105
 1318 TCGTGGCCCAAGGCCACCTTGGGATCGATTAA 1349
 2106 CCGTGGCCCAAGGCCACCTTGGGGCCCCCATTA 2137
 RESULT 8
 AAV52933
 ID AAV52933 standard; cDNA; 2537 BP.
 XX AAV52933;
 XX AC
 DT 25-MAR-2003 (updated)
 DT 21-DEC-1998 (first entry)
 XX Human pre-transforming growth factor-beta 1 cDNA.
 DE Transforming growth factor-beta 1; TGF-beta 1; human; ss.
 KW Transforming growth factor-beta 1; TGF-beta 1; human; ss.
 XX Homo sapiens.
 OS
 FH Location/Qualifiers
 FT CDS 842..2014
 FT /tag= a
 FT mat_peptide 1676..2011
 FT /tag= b
 FT stem_loop 37..113
 FT /tag= b
 FT /note= "putative stable hairpin loop"
 FT misc_feature 2015..2100
 FT /tag= c
 FT /note= "GC-rich sequence"
 FT polyA_signal 2514..2520
 FT /tag= d
 XX US5801231-A.
 XX 01-SEP-1998.
 XX 30-MAY-1995; 95US-0454468.
 XX 13-MAR-1987; 87US-0025423.
 XX 22-MAR-1985; 85US-0715142.
 XX 04-AUG-1989; 89US-0389929.
 XX 04-MAR-1992; 92US-0845893.
 XX 05-NOV-1993; 93US-0147364.
 XX 30-MAY-1995; 95US-0454468.
 XX (GETH) GENENTECH INC.
 XX Derynck RNA, Goeddel DV;
 XX WPI: 1998-494840/42.
 XX P-PSDB; AAW78785.
 XX DNA encoding transforming growth factor-beta precursor sequence -
 useful for analysis to perform manipulations to increase yield of
 recombinant production of the protein
 XX Example 3; Fig 1B 1-3; 26pp; English.
 XX This nucleotide sequence codes for the human transforming growth
 factor-beta 1 precursor (preTGF-beta 1, see AAW78785). It is a
 composite of overlapping cDNA clones isolated from different cDNA
 libraries (placenta, A172 glioblastoma, HT1080 fibroblastoma) using
 TGF-beta exon (see AAV52933) restriction fragments as probes.
 CC The 3' region of the sequence was determined using cloned genomic
 DNA. The invention relates to the recombinant production of

[illegible]

RESULT 13
AAN81084
ID AAN8
XX
AC AAN8
XX

501	TACAACAGCAGCCCGCCAGCCGGTGGCCGGGAGAGTGGCGAGCCGGACCCCAACCGGAG	560
311	CGCGACTACTACGCCAAGGAGGTGACCCCGCGTGTAAATGGTGGAAAGCGCAACCAAAATC	370
561	CGCCACTACTACGCCAAGGAGGTGACCCCGCGTGTAAATGGTGGAAAGCGCAACCAAAATC	620
371	TAICATAAAATTCAGGGCAGCCGCCACAGCTTATATATGCTGTTCACACAGTGGAGCTC	430
621	TATGACAAAGTTCAAGCAGAGCAGACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
431	CGGAAGCGGTGGCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCTGCTGAGGCTC	490
681	CGAAGCAGTACTGTAACCTGTGTGCTCTCCGGCAGAGCTGCGTCTGCTGAGGCTC	740
491	AAGTTAAAGTGGAGCAGCAGCTGGAGCTATTACACAGAAATACACGAATGATTCTTGCGGC	550
741	AAGTTAAAGTGGAGCAGCAGCTGGAGCTGTACCAAGAAATACAGCAAAATCTCTGGCGA	800
551	TACCTCAGCAACCGGCTGCTGGCCGCCAGTACTCACCGGAGTGGCTGTCTTTGATGTC	610
801	TACCTCAGCAACCGGCTGCTGGCCGCCAGCAACTCGCGGAGTGGTGTCTTTTATGATC	860
611	ACCGGAGTTGTGGCAGTGGCTGACCCCGCAGAGAGGCTATAGAGGTTTTTCGCCTCAGT	670
861	ACCGGAGTTGTGGCAGTGGTGTGAGCCGCGAGGGGAAATGAGGGCTTTTCGCCTTAGC	920
671	GCCCACTCTTCTCTGACAGCAAGATACACACTCCAGTGGGAAATTAAGCGGTTCAAT	730
921	GCCCACTCTCTCTGTGACAGCAAGATACACACTGCAAGTGGACATCAACGGTTCACT	980
731	TCTGGCCCGGGGTGACCTGGCCAGCATTCACGGGATGACCGGCCCTTCTCTGCTCTC	790
981	ACCGCCGCCAGTGGACCTGGCCAGAAATTCATGGCATGAACGGCCCTTCTCTGCTCTC	1040
791	ATGGCCACCCCGTGGAGAGGGCCAGCACTGTCACAGCTCCCGGCACCGCGAGCCCTG	850
1041	ATGGCCACCCCGTGGAGAGGGGCCAACATCTGCAAGCTCCCGGCACCGCGGAG----	1095
851	GATACCAACAGCTACCATAAGCATGCGAGTGGCAGACTACGCATCTCTGGCCCTGGATACCAAC	910
1096	-----T-----CCCTGGACACCAAC	1109
911	TACTGCTTACGCTCCAGGAGAACTGCTGCTGCGCAGCTACATGACTTCACTTCCGG	970
1110	TACTGCTTACGCTCCAGGAGAACTGCTGCTGCGCAGCTGTATATGACTTCCGC	1169
971	AAGGACCTGGCTGGAATGATTCATGAACCAAGGGTACCATGCGCAATTTCTGCGTG	1030
1170	AAGGACCTGGCTGGAATGATTCACAGAGCCCAAGGGTACCATGCGCAATTTCTGCGTG	1229
1031	GGGCGCTGTCCTTACATCTGGAGCCTAGACNCTCAGTACAGCAAGTCTTGGCTGTGAC	1090
1230	GGGCGCTGTCCTTACATTTGGAGCCTGGACAGCAGTACAGCAAGTCTTGGCGCTGTAC	1289
1091	AACAGCACAACCCGGCGGCTCGCGGCCCTGCTGCTGCGCAGCGCTGGAGCCA	1150
1290	AACAGCATAACCCGGCGGCTCGCGGCCCTGCTGCTGCGCAGCGCTGGAGCCA	1349
1151	CTGCCCATCTGTACTAGTGGGCGCGCAGCCCAAGGTGGAGCAGTGTCCACATGATC	1210
1350	CTGCCCATCTGTACTAGTGGGCGCAAGCCCAAGGTGGAGCAGTGTCTCAACATGATC	1409
1211	GTGCGCTTCTTCAAGTGCAGCTGAGGCCCGCCGCCACAGCCCGCCACCGCGGAG	1270
1410	GTGCGCTTCTTCAAGTGCAGCTGAGGCCCGCCGCCACAGCCCGCCACCGCGGAG	1469
1271	GCCCGGCCCAACCCCGCCGCGCT-----CACCGGGGCTGTATTATTAAGGACA-TGCT	1321
1470	GCCCGGCCCAACCCCGCGCTGTCTGCTGCTGGGGCTGTATTTAAGCACACCGGT	1529
1322	GCCCAAGCCCACTGGGATCGATTAA	1349
1530	GCCCAAGCCCACTGGGCGCCATTAA	1557


```
Db 1317 GCGCTCAAGTAAAGTGGACGAGCAGCTGGAGCTGTACCAAGAAATACAGCAAAATTCCT 1376
QY 546 GCGCTAGCTCAGCAACCGCTGTGTGGCCCGCAGTGACTACCGGAGTGGTGTCTTTG 605
Db 1377 GCGCATACCTCAGCAACCGCTGTGTGGCCACCGGACTCGCCAGAGTGTATCTTTG 1436
QY 606 ATGTACCGGAGTGTGGCCGAGTGGCTGAGCCGACAGAGGCTATAGAGGTTTTCGCC 665
Db 1437 ATGTACCGGAGTGTGGCCGAGTGGCTGAGCCGAGGAGGAAATAGAGGCTTTTCGCC 1496
QY 666 TCAGTGGCCACTTTCCTCTGACAGCAAGATACACACTCCAGCTGAAATTAACGGT 725
Db 1497 TTAGCGCCACTGCTCTGTGACAGCGGATACACACTCCAGTGGACATCAACGGT 1556
QY 726 TCAATTTGGCCGCGGGGTGACCTGGCCACCATTCAGGGATGAACGGGCCCTTCCTGC 785
Db 1557 TCATACCGGCGCGGAGTGAACCTGGCCACCATTCATGGCATGAACGGGCTTTCCTGC 1616
QY 786 TCCTCATGGCCACCGCTGGAGAGGCGCCAGCAGCTCCAGAGTCCCGGACCGCCGAG 845
Db 1617 TTCTCATGGCCACCGCTGGAGAGGCGCCAGCATCTGCAAGCTCCCGGACCGCGGA- 1675
QY 846 CCTCATACCAACAGCTACCATACGACGTCGCCAGCTACGATCTCTGCGCCCTGGATA 905
Db 1676 -----GCCCTGGACA 1685
QY 906 CCAACTACTGTTTACGCTCCAGGAGAGAACTGCTGCGCGGAGCTCTACATTGACT 965
Db 1686 CCAACTATGTTTACGCTCCAGGAGAGAACTGCTGCGCGGAGCTCTACATTGACT 1745
QY 966 TCCGGAAGGACTGCGGCTGGAAGTGGATTGATGAACCAAGGGGTACCATGCCAATTC 1025
Db 1746 TCCGCAAGGACTGCGGCTGGAAGTGGATCCAGAGCCCAAGGGCTACCATGCCAATTC 1805
QY 1026 GCGTGGGCGCTGCTTACATCTGAGGCTAGACACTAGTACAGCAAGTCTCTGGCTC 1085
Db 1806 GCGTGGGCGCTGCTTACATCTGAGGCTAGACACTAGTACAGCAAGTCTCTGGCTC 1865
QY 1086 TGTACACCAAGCAGCAACCGCGCGCTGCGGCGCGCTGCTGCTGCGCGGAGGCTGG 1145
Db 1866 TGTACACCAAGCAGCAACCGCGCGCTGCGGCGCGCTGCTGCTGCGCGGAGGCTGG 1925
QY 1146 AGCCACTGCCCCCTGTACTAGTGGCGCGCAAGCCCAAGGTGAGAGAGTGTCCCAACA 1205
Db 1926 AGCCCTGCCCCCTGTACTAGTGGCGCGCAAGCCCAAGGTGAGAGAGTGTCCCAACA 1985
QY 1206 TGATGTGGTCTCTGCAAGTGCAGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1265
Db 1986 TGATGTGGTCTCTGCAAGTGCAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2045
QY 1266 GGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
Db 2046 GGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2105
QY 1318 TGTGTCGCCCAAGCCACTGTG3ATCGATTAA 1349
Db 2106 CCGTGGCCCAAGCCACTGTG3GCGCCCAATTA 2137
```

RESULT 2

```
US-07-669-171-1
; Sequence 1, Application US/07669171
; Patent No. 5304541
; GENERAL INFORMATION:
; APPLICANT: PURCHIO, ANTHONY P.
; APPLICANT: MADISEN, LINDA
; APPLICANT: MERWIN, JUNE RAE
; TITLE OF INVENTION: TGF- $\beta$ 1/2: A NOVEL CHIMERIC TRANSFORMING
; TITLE OF INVENTION: GROWTH FACTOR-BETA
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
```

```
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/669,171
; FILING DATE: 19910314
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-159-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 261..1430
; US-07-669-171-1
```

Query Match 71.3%; Score 970.8; DB 1: Length 1560;

Best Local Similarity 85.0%; Pred. No. 1.6e-203;

Matches 1146; Conservative 0; Mismatches 142; Indels 60; Gaps 3;

```
QY 11 ATGCGCGCTTCGGGGTGGGGCTCTTGGCGCTGCTGCTGCGCTGTGTGGTGTAGT 70
Db 261 ATGCGCGCTTCGGGGTGGGGCTCTTGGCGCTGCTGCTGCGCTGTGTGGTGTAGT 320
QY 71 CTGACCGCTTGGCGCGCGCGCGCGCGCTGCTGCTGCAACCTGCAACACCATGACATG 130
Db 321 CTGACCGCTTGGCGCGCGCGCGCGCGCTGCTGCTGCAACCTGCAACACCATGACATG 380
QY 131 GTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAAGCTTCGGCT 190
Db 381 GTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAAGCTTCGGCT 440
QY 191 AGCCCGCGGAGCGGAGCGGAGCTGCGCGCGCGCGCGCGCTGCTGAGGAGTGTGCTT 250
Db 441 AGCCCGCGGAGCGGAGCGGAGCTGCGCGCGCGCGCGCGCTGCTGAGGAGTGTGCTT 500
QY 251 TACAACAGTACCGCGCGCGCGGTTAGCGGGGAAAGTGTGCAACCGGAGCCCGAGCAGAG 310
Db 501 TACAACAGTACCGCGCGCGCGGTTAGCGGGGAAAGTGTGCAACCGGAGCCCGAGCAGAG 560
QY 311 GCGGACTTACTACGCGCAAGGAGTCAACCGCGCTGCTTAATGTGGAAGCGGACCAAAATC 370
Db 561 GCGGACTTACTACGCGCAAGGAGTCAACCGCGCTGCTTAATGTGGAAGCGGACCAAAATC 620
QY 371 TATGATAAATTAAGGCGCACCCCGGAGCTATATATGCTGTTCACACGTCGGAGCTC 430
Db 621 TATGATAAATTAAGGCGCACCCCGGAGCTATATATGCTGTTCACACGTCGGAGCTC 680
QY 431 CGGAAGCGGTGCGGGAACCTGTTGCTCTCTGGGAGAGCTGCGGCTGCTGAGGCTC 490
Db 681 CGGAAGCGGTGCGGGAACCTGTTGCTCTCTGGGAGAGCTGCGGCTGCTGAGGCTC 740
QY 491 AAGTTAAAGTGGAGCAGCGTGGAGCTATACCAAGAAATACAGCAAAATTCCTTGGCG 550
Db 741 AAGTTAAAGTGGAGCAGCGTGGAGCTATACCAAGAAATACAGCAAAATTCCTTGGCG 800
```



```

Db 1437 ATGTCACCGGAGTGTGCGGGAGTGTTGAGCGTGGAGGGGAAATTAGGGCTTTTCGCC 1496
Qy 666 TCAGTGGCCCACTTCTCCTTSACAGCAAGATTAACACACTCCACGCTGGAAATTAACGGGT 725
Db 1497 TTABCGCCCACTGCTCCTGTGSACAGCAGGATTAACACACTGCAAGTGGACATCAACGGGT 1556
Qy 726 TCAATTCGTGGCCCGGGGTGACTGCGCACCAATTCACGGCATGAACGGCCCTTCTGCG 785
Db 1557 TCATAGCGCGCGGAGGTGACTGCGCACCAATTCATGGCATGAACGGCCCTTCTGCG 1616
Qy 786 TCCTCATGGCCACCCCGCTGAGAGGCGCCAGACACTGACACACTCCCGCACCGCGGAG 845
Db 1617 TTCTCATGGCCACCCCGCTGAGAGGCGCCAGCATCTGCAAACTCCCGCACCGCGGAG 1675
Qy 846 CCCTGGATACACAGCTACCCATACAGCTGCGACACTACGCACTCTCTGGGCCCTGGATA 905
Db 1676 -----GCCCTGGACA 1685
Qy 906 CCAACTACTGCTTCAGCTCCACGGAGAAAGTCTGCTGCGCGGAGCTCTACATTGACT 965
Db 1686 CCAACTATTGCTTCAGCTCCACGGAGAAAGTCTGCTGCGCGGAGCTGTACATTGACT 1745
Qy 966 TCCGGAAGGAGCTGGGTGGAGTGGATTTCATGAACCCAAAGGCTACCATGCAATTTCT 1025
Db 1746 TCCGGAAGGAGCTGGGTGGAGTGGATTCCACGAGCCCAAGGCTACCATGCAATTTCT 1805
Qy 1026 GCCTGGGGCGCTGCTTACATCTGAGCGCTAGACACTAGTACAGCAAGTCTCTGCTC 1085
Db 1806 GCCTGGGGCGCTGCTTACATCTGAGCGCTGAGCGCGCTAGACCGAGTACAGCAAGTCTCTGCTC 1865
Qy 1086 TGTAAACACAGCAACCCCGGCGCTGCGGGCGCGTCTGCTGCGCGGAGCGCTGG 1145
Db 1866 TGTAAACACAGCAATAACCGGCGCTGCGGGCGCGTCTGCTGCGCGGAGCGCTGG 1925
Qy 1146 AGCCACTGCCATCGTGTAC:ACGTGGCGCGCAAGCCCAAGTGGAGCAGCTGTCCAACA 1205
Db 1926 AGCCGCTGCCATCGTGTAC:ACGTGGCGCGCAAGCCCAAGTGGAGCAGCTGTCCAACA 1985
Qy 1206 TGATCGTGGTTCCTGCAAG:GCAAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1265
Db 1986 TGATCGTGGTTCCTGCAAG:GCAAGTGGAGTGGTCCCGCGCGCGCGCGCGCGCGCGCG 2045
Qy 1266 GCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
Db 2046 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2105
Qy 1318 -TCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1349
Db 2106 CCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2138

```

RESULT 4

```

US-09-661-753-28
; Sequence 28, Application US/09661753
; Patent No. 6436909
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan F. Murray
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR BETA
; FILE REFERENCE: ISPU-0498
; CURRENT APPLICATION NUMBER: US/09/661,753
; EARLIER APPLICATION DATE: 2000-09-14
; EARLIER FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 68
; SEQ ID NO 28
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (842)...(2017)
US-09-661-753-28

```

```

Query Match 71.0%; Score 965.8; DB 4; Length 2745;
Best Local Similarity 84.6%; Pred. No. 2.2e-202;
Matches 1146; Conservative 0; Mismatches 147; Indels 62; Gaps 3;

Qy 6 CCGAGATGGCGGCTTCGGGGCTGGGCTCTTGGCGCTGCTGCTGCGCTGCTGCTGCTGCTG 65
Db 837 CCCCCATCCCGGCTCCCGGCTGGGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896
Qy 66 TAGTCTGACGCTGCGCGCGCGCGCGCGGACTGCTCACCTGCAAGACCATCAACATGG 125
Db 897 TGGTCTGACGCTGCGCGCGCGCGCGGACTATCCACCTGCAAGACTATCGACATGG 956
Qy 126 AGCTGGTGAAGGGAGGCGCATCGAGGCCATTCGGGGCCAGATTCTGTCCAAAGCTTGGC 185
Db 957 AGCTGGTGAAGGGAGGCGCATCGAGGCCATTCGGGGCCAGATTCTGTCCAAAGCTTGGC 1016
Qy 186 TTGCCAGCCCCCGAGCGGGGAGCTGCGCGCGCGCGCGCGCTGCTGCTGAGGAGTACTGG 245
Db 1017 TCGCCAGCCCCCGAGCGGGGAGGTGCGCGCGCGCGCGCGCTGCTGCTGAGGCGCTGCTG 1076
Qy 246 CTCCTTACACAGTACCGCGCGCGCGGAGTGGCGGGGAAAGTGTGCAACCGGAGCCGAGC 305
Db 1077 CCTCTTACACAGTACCGCGCGCGCGGAGTGGCGGGGAGTGCAGAACCGGAGCCGAGC 1136
Qy 306 CAGAGCGGAGCTACTACGCCAAGGAGTCAACCGCGTCTAATGTTGGAAGCGGCAACC 365
Db 1137 CTGAGCGGAGCTACTACGCCAAGGAGTCAACCGCGTCTAATGTTGGAAGCGGCAACC 1196
Qy 366 AAATCTATGATTAATTCGAAGGCGACCGCGCGGAGTGTATATATATATGCTGTTCACAGCTGG 425
Db 1197 AAATCTATGATTAATTCGAAGGCGAGTACACAGCATATATATGTTCTTCAACATCAG 1256
Qy 426 AGCTCGGGAAGCGGTGCGCGGAACTGTTCTCTCTCGGCGAGCTGCGCTGCTGCTGCTG 483
Db 1257 AGCTCGGGAAGCGGTGCTGAAACCGTGTCTCTCTCGGCGAGCTGCGCTGCTGCTGCTG 1315
Qy 484 -GAGGCTCAAGTTAAAGTGGAGCAGCAGTGGAGCTATACAGAAATACAGCAATGATT 542
Db 1317 GGAGGCTCAAGTTAAAGTGGAGCAGCAGTGGAGCTGTACCAAGAAATACAGCAATGATT 1376
Qy 543 CCGTGGGCTTACCTGAGCAACCGGCTGCTGGCGCGCGCGGAGTCACTGCGGAGTGGCTGCT 602
Db 1377 CCGTGGGCTTACCTGAGCAACCGGCTGCTGGCAGCGGCTGCGGAGGCTGCGGAGTGGT 1436
Qy 603 TTGATGTCAACCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGGAGGCTATAGAGGTTTC 662
Db 1437 TTGATGTCAACCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGGAGTGTGAGGAGTTC 1496
Qy 663 GCGTCAAGTGGCGGCTTCTCTCTGACAGCAAGATTAACACACTTCCAGCTGGAATTAAG 722
Db 1497 GCGTCAAGTGGCGGCTTCTCTCTGACAGCAAGATTAACACACTTCCAGTGGAGTCAAG 1556
Qy 723 GGTTCATTTCTGGCGCGCGGCTGACCTGGCCACCATTCACGGCATGAACCGGCTTCC 782
Db 1557 GGTTCATTTCTGGCGCGCGGCTGACCTGGCCACCATTCATGCGATGAACCGGCTTCC 1616
Qy 783 TGCTCTCATGCCACCGCGCTGGAGAGGCGCGGAGCTGACAGCTGCGGCGGAGCGGCG 842
Db 1617 TGCTCTCATGCCACCGCGCTGGAGAGGCGCGGAGCTGCAAGCTCCCGGAGCGGCG 1676
Qy 843 GAGCCCTGGATACCAACAGCTACCCATAGGAGTGGCAGACTACGCATCTCTGCGGCTGG 902
Db 1677 GA-----GCGCTGG 1685
Qy 903 ATACCAACTACTGCTTACGCTCCACGGAGAAAGTCTGCTGCGGCGGAGCTCTACATTG 962
Db 1686 ACACCAACTACTGCTTACGCTCCACGGAGAAAGTCTGCTGCGGCGGAGCTCTACATTG 1745
Qy 963 ACTTCCGGAAGGAGCTGGGCTGAAAGTGCATTAACCAAGGCTACCATGCCAATT 1022
Db 1746 ACTTCCGGAAGGAGCTGGGCTGAAAGTGCATTAACCAAGGCTACCATGCCAATT 1805

```


	Query Match	56.8%	Score 773;	DB 6;	length 2207;
	Best Local Similarity	78.4%	Pred. No. 3e-160;		
	Matches	963;	Conservative	0;	Mismatches 215; Indels 51; Gaps 1;
Qy	6	CCGAGATGGCGGCCTTCGGGGCTCGGGCTCTTCGCGCTGCTGCTGCCGCTGCTGTGGGTGC	65		
Db	256				
Qy	66	CCCCATGCCGCGCCTCCGGCTCGGGCTGTGCGCTGCTGCTACCGCTCTGTGGGTAC	315		
Db	316	TAGTGCTGACGCTTCGGCGCGCGCCGCGGACTGTCCACCTGCAACACATCGACATGG	125		
Qy	126	AGCTGTTGAAGCGAAGCGCATCAGAGCCCATTCGCGGCCAGATTCTCTCAAAGTCTCGGC	185		
Db	376	AGCTGTTGAAGCGAAGCGCATCAGAGCCATCCGCGGCCAGATCTGTCCAAGCTCGGC	435		
Qy	186	TTGCCAGCCCCCGAGCAGGGGACGTGCGCCCGGCCGCTGCCTGAGACGTACTATGG	245		
Db	436	TCGCCAGCCCCCGAGCAGGGGAGTTCGCGCCCGGCCGCTGCCGAGACCGTGTCTCG	495		
Qy	246	CTCTTTTACAACAGTACCCGGACCGGGTAGCCGGGGAAAGTGTTCGAACCGGAGCCCGAGC	305		
Db	496	CCCTGTACAACAGCACCGCGACCCGGTGGCGGGGAGAGCGCGGAGCCGGAGCCCGAAC	555		

```

QY 306 CAGAGCGGAGTACTAGGC3AGGAGGTGACCCCGCTGCTAATGCTGGAAAGCGGCAACC 365
Db 556 CGGAGCGGAGTACTAGGC3AGGAGGTGACCCCGCTGCTAATGCTGGAAAGCGGCAACC 615
QY 366 AATCTATGATAAATCAA3GGGACCCGCCACAGCTTATATATGCTTCAACACGCTGG 425
Db 616 AATCTATGACAAGTCAA3CAGAGCACACACAGCATATATATGCTTCAACACATCAG 675
QY 426 AGCTCGGGAGCGGTGCC3GAACCTGTATGCTCTCTCGGGCAGAGCTGGCGCTCTGA 485
Db 676 AGCTCGGAGAAGCAGTACC3GAACCTGTGTGCTCTCCCGGCAGAGCTGGCTCTGCTGA 735
QY 486 GGCTCAAGTAAAGTGA3CAGCAGGTGGAGCTATACCAAGAAATACAGCAATGATTCCT 545
Db 736 GGCTCAAGTAAAGTGA3CAGCAGTGGAGCTGTACCAAGAAATACAGCAATTCCT 795
QY 546 GCGCTACCTCAGCAACCG3CTGCTGCCCGCCAGTGACTCACCGGAGTGCTGCTCTTTG 605
Db 796 GCGCTACCTCAGCAACCG3CTGCTGCCCGCCAGCACTCCCGGAGTGGTGTCTTTG 855
QY 606 ATGTCACCGGAGTGTGCG3AGTGGCTGACCCGCGAGAGGCTATAGAGGGTTTTGCGC 665
Db 856 ATGTCACCGGAGTGTGCG3AGTGGTGTAGCGCGGAGGAAATGAGGGCTTTGCGC 915
QY 666 TCAGTGCCCACTCTTCTCTGACAGCAAGATACACACTCCAGCTGGAAATTAACGGGT 725
Db 916 TTAGCGCCCACTGCTCTGTGACAGCAAGATACACACTGCAAGTGAGCATCAACGGGT 975
QY 726 TCAATTCGCGCGCGGGTGACTGCGCCACCATTCACGGCATGAACCGCGCTTCTCTG 785
Db 976 TCACTACCGCGCGGAGTGAGCTGCGCCACAATTCATGGCATGAACCGCGCTTCTCTG 1035
QY 786 TCCTCATGGCCACCCGCT3GAGAGGGCCAGCACTGCACAGCTCCCGCAGCGCGGAG 845
Db 1036 TTCTCATGGCCACCCCACT3GAGAGGGCCCAACAICTGCAAGCTCCCGCAGCGCGGA - 1094
QY 846 CCCTGGATACCAACAGCTA3CCATACAGCTGCCAGACTACGCATCTCTGGCCCTGGATA 905
Db 1095 -----GCITGGATC 1104
QY 906 CCAACTACTGTTGAGCTC3ACGGAGAAAGTACTGCTGGTGGCGAGCTCTACATGACT 965
Db 1105 CGGCTATGCTTTAGAAA1GTCCAGCATAA1TGTCTGCTTACCTCCACTTTACATGATT 1164
QY 966 TCCGGAAGGACCTGGGCTG3AAGTGGATTGATGAACCCAGGCTACCATGCCAAATTTCT 1025
Db 1165 TCAAGAGGATCTAGGTG3AAATGGATACAGNACCCAAAGGTAGCAATGCCAACTTCT 1224
QY 1026 GCCTGGGCGCTCTCCCTA3ATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTC 1085
Db 1225 GTGCTGGAGCATGCCGTA1TTATGGAGTTCAGACACTCAGCACAGCGGCTCTGAGCT 1284
QY 1086 TGTACACACGACACAACC3GGCGCGTGGGGCGCGCTGCTGGTGGCGAGCGGCTGG 1145
Db 1285 TATATAATACCATAAAT3CAGAGCATCTGCTTCTCTGCTGGTGTCCCAAGATTAG 1344
QY 1146 AGCCACTGCCCATCGTGTACTAGCTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1205
Db 1345 AACCTTAACCATTTCTCTA3TACATTTGGCAAAACACCCCAAGATTGAACAGCTTTCTAATA 1404
QY 1206 TGATGCTGCTGCTCTCTGCAAT3TGCAAGCTGA 1234
Db 1405 TGATGTAAGTCTTTGCAATGCACTAA 1433

```

RESULT 10

US-08-470-837-29
: Sequence 29, Application US/08470837
: Patent No. 5800811
: GENERAL INFORMATION:
: APPLICANT: Nimni, Marcel F.
: APPLICANT: Hall, Frederick L.

```

: APPLICANT: Tuan, Tai-Ian
: APPLICANT: Wu, Lingtao
: APPLICANT: Cheung, David T.
: TITLE OF INVENTION: Transforming Growth Factor B Fusion
: TITLE OF INVENTION: and
: TITLE OF INVENTION: Their Use in Wound Healing
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merchant & Gould
: STREET: 11150 Santa Monica Boulevard, Suite 400
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90025-3395
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/470,837
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharp, Janice A.
: REGISTRATION NUMBER: 34,051
: REFERENCE/DOCKET NUMBER: 30630-1US01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 310-445-1140
: TELEFAX: 310-445-9031
: INFORMATION FOR SEQ ID NO: 29:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 339 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..333
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..336
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 1
:
US-08-470-837-29

Query Match 22.8%; Score 310.2; DB 1; Length 339;
Rest Local Similarity 94.7%; Pred. No. 3.1e-59;
Matches 321; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 896 GCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAGAACTGCTGCTGGCGGAGCTC 955
Db 1 GCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAACTGCTGCTGGCGGAGCTG 60
QY 956 TACATTGACTTCCGGAAGCACCTGGCTGGAAGTGATTCATGACCAACCAAGGGCTACCAT 1015
Db 61 TACATTGACTTCCGGAAGCACCTGGCTGGAAGTGATTCATGACCAACCAAGGGCTACCAT 120
QY 1016 GCAATTTCTGCTGGGCGCTGTCCCTACATCTGAGCCTAGACACTAGACAGCAAG 1075
Db 121 GCAACTTCTGCTGGGCGCTGTCCCTACATTTGGAGCTGGACAGCAGTACAGCAAG 180
QY 1076 GTCTGGCTCTGTACAAACAGCACACCCGGCGGCTGGCGGCGCGCTGCTGCTGGCG 1135
Db 181 GTCTGGCGCTGTACAAACAGCATAAACCGGCGCGCTGGCGGCGCGCTGCTGCTGGCG 240
QY 1136 CAGGCGCTGGAGCCACTGCCATCGTGTACTAGTGGCGCGCAAGCCCAAGGTGGAGCAG 1195
Db 241 CAGGCGCTGGAGCCGCTGCCCATCGTGTGTACTAGTGGCGCGCAAGCCCAAGGTGGAGCAG 300
QY 1196 CTGTCCAACATGATCGTGGTTCCTCTCAAGTGCAGCTGA 1234

```

DB 301 CTGTCCAAACATGATCGTGGCTCTCTGCAAGTGCAGCTGA 339
|||||

RESULT 11

US-08-486-057B-1
: Sequence 1, Application US/08486057B
: Patent No. 5650494
: GENERAL INFORMATION:
: APPLICANT: Cerletti, Nico
: APPLICANT: McMaster, Gary K.
: APPLICANT: Cox, David
: APPLICANT: Schmitz, Albert
: APPLICANT: Meyhack, Bernd
: TITLE OF INVENTION: Process for Refolding Recombinantly
: TITLE OF INVENTION: Produced TGF-beta-like Proteins
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Henry P. No. 5650494ak
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10591-9005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: APPLICATION DATA:
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/201,703
: FILING DATE: 25-FEB-1994
: APPLICATION DATA:
: APPLICATION NUMBER: US 07/960,309
: FILING DATE: 13-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/521,502
: FILING DATE: 03-DEC-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 8927546.5
: FILING DATE: 06-DEC-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5650494ak, Henry P.
: REGISTRATION NUMBER: 33200
: REFERENCE/DOCKET NUMBER: 4-17861/+ /Cont3
: TELEPHONE: (908) 277-5110
: TELEFAX: (908) 277-4306
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 339 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..339
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..339
US-08-486-057B-1

Query Match 22.7%; Score 308.6; DB 1; Length 339;
Best Local Similarity 94.4%; Ref. No. 6.9e-59;
Matches 320; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
956 GCCTGGAGCAACTATTCTTCCAGCTCCACGGAGAGAACTGCTGGCTGGGCGAGCTG 60
1015
61 TACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCAT 120
1016 GCCAAATTTCTGCTGGGCGCCCTGCTTACATCTCTGAGCCTAGACACTCAGTACAGCAAG 1075
180
121 GCCAACTTCTGCTGGGCGCCCTGCCCCCTACATTTGGAGCCTGGACACGAGTACAGCAAG 180
1135
1076 GTCTGGGCTCTGTACAACAGCACAAACCGGGCGGCTCGGGGGCGCCGCTGCTGGTGGCG 1135
240
181 GTCTGGGCTCTGTACAACAGCACAAACCGGGCGGCTCGGGGGCGCCGCTGCTGGTGGCG 240
1136 CAGGCGCTGGAGCCACTGCCCATCTGTACTACGTGGGCGCCGAGCCCAAGGTGGAGCAG 1195
241 CAGGCGCTGGAGCCGCTGCCCATCTGTACTACGTGGGCGCCGAGCCCAAGGTGGAGCAG 300
1196 CTGTCCAAACATGATCGTGGCTCTCTTCCCTGCAAGTGCAGCTGA 1234
301 CTGTCCAAACATGATCGTGGCTCTCTTCCCTGCAAGTGCAGCTGA 339

RESULT 12

US-08-789-588-1
: Sequence 1, Application US/08789588
: Patent No. 5922846
: GENERAL INFORMATION:
: APPLICANT: Cerletti, Nico
: APPLICANT: McMaster, Gary K.
: APPLICANT: Cox, David
: APPLICANT: Schmitz, Albert
: APPLICANT: Meyhack, Bernd
: TITLE OF INVENTION: Process for Refolding Recombinantly
: TITLE OF INVENTION: Produced TGF-beta-like Proteins
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Henry P. No. 5922846ak
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10591-9005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/789,588
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/486,057
: FILING DATE: 07-JUN-1995
: APPLICATION NUMBER: US 08/201,703
: FILING DATE: 25-FEB-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/960,309
: FILING DATE: 13-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/621,502
: FILING DATE: 03-DEC-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 8927546.5
: FILING DATE: 06-DEC-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5922846ak, Henry P.
: REGISTRATION NUMBER: 33200
: REFERENCE/DOCKET NUMBER: 4-17861/+ /Cont3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 277-5110
: TELEFAX: (908) 277-4306

```
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 339 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..339
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..339
; US-08-789-588-1

Query Match      22.7%; Score 308.6; DB 2; Length 339;
Best Local Similarity 94.4%; Pred. No. 6.9e-59;
Matches 320; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      896  GCCTGGATACCAACTACTGCTTCCAGCTCCACGGAGAAGAACTGCTGCGTGGGCGAGCTC 955
Db      1      GCCTGGACACCAACTATTGCTTCCAGCTCCACGGAGAAGAACTGCTGCGTGGGCGAGCTG 60

QY      956  TACATTGACTTCCCGAAGGACCTGGCTTGGAACTGATTTCATGAACCCAAAGGCTACCAT 1015
Db      61      TACATTGACTTCCCGAAGGACCTGGCTTGGAACTGATTTCATGAACCCAAAGGCTACCAT 120

QY      1016  GCCAATTTCCTGCTGGGCGCTGTCCTTACATCTGAGCCTAGACACTCACTACAGCAAG 1075
Db      121     GCCAATTTCCTGCTGGGCGCTGTCCTTACATCTGAGCCTAGACACTCACTACAGCAAG 180

QY      1076  GTCTGGCTGTGTACACAGCACAAACCGGGCGCTCGGGCGGCGCTGCTGCGTGGCG 1135
Db      181     GTCTGGCGCTGTGTACACAGCATAAACCGGGCGCTCGGGCGGCGCTGCTGCGTGGCG 240

QY      1136  CAGGCGCTGGAGCCACTGCCCATCTGTACTAGTGGGCGCGCAAGCCCAAGTGGAGCAG 1195
Db      241     CAGGCGCTGGAGCCCTGCCCATCTGTACTAGTGGGCGCGCAAGCCCAAGTGGAGCAG 300

QY      1196  CTGTCCAACATGATGCTGCTTCTCTCAAGTGCAGCTGA 1234
Db      301     CTGTCCAACATGATGCTGCTCTCTCAAGTGCAGCTGA 339

RESULT 14
US-09-169-768-9
; Sequence 9, Application US/09169768
; Patent No. 6492508
; GENERAL INFORMATION:
; APPLICANT: GRUSKIN, ELLIOT A.
; APPLICANT: BUECHTER, DOUGLAS
; APPLICANT: BROKAW, JANE
; APPLICANT: ZHANG, GUANGHUI
; APPLICANT: PAOLELLA, DAVID
; TITLE OF INVENTION: AMINO ACID MODIFIED POLYPEPTIDES
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DILLWORTH & BARRESE
; STREET: 333 EARLE OVINGTON BOULEVARD
; CITY: UNIONDALE
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 11553
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/169,768
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: STEEN, JEFFREY S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 228-8484
; TELEFAX: (516) 228-8516
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;

;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 339 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..339
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..339
; US-08-789-588-1

Query Match      22.7%; Score 308.6; DB 2; Length 339;
Best Local Similarity 94.4%; Pred. No. 6.9e-59;
Matches 320; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      896  GCCTGGATACCAACTACTGCTTCCAGCTCCACGGAGAAGAACTGCTGCGTGGGCGAGCTC 955
Db      1      GCCTGGACACCAACTATTGCTTCCAGCTCCACGGAGAAGAACTGCTGCGTGGGCGAGCTG 60

QY      956  TACATTGACTTCCCGAAGGACCTGGCTTGGAACTGATTTCATGAACCCAAAGGCTACCAT 1015
Db      61      TACATTGACTTCCCGAAGGACCTGGCTTGGAACTGATTTCATGAACCCAAAGGCTACCAT 120

QY      1016  GCCAATTTCCTGCTGGGCGCTGTCCTTACATCTGAGCCTAGACACTCACTACAGCAAG 1075
Db      121     GCCAATTTCCTGCTGGGCGCTGTCCTTACATCTGAGCCTAGACACTCACTACAGCAAG 180

QY      1076  GTCTGGCTGTGTACACAGCACAAACCGGGCGCTCGGGCGGCGCTGCTGCGTGGCG 1135
Db      181     GTCTGGCGCTGTGTACACAGCATAAACCGGGCGCTCGGGCGGCGCTGCTGCGTGGCG 240

QY      1136  CAGGCGCTGGAGCCACTGCCCATCTGTACTAGTGGGCGCGCAAGCCCAAGTGGAGCAG 1195
Db      241     CAGGCGCTGGAGCCCTGCCCATCTGTACTAGTGGGCGCGCAAGCCCAAGTGGAGCAG 300

QY      1196  CTGTCCAACATGATGCTGCTTCTCTCAAGTGCAGCTGA 1234
Db      301     CTGTCCAACATGATGCTGCTCTCTCAAGTGCAGCTGA 339

RESULT 13
US-123-233-1
; Sequence 1, Application US/0912333
; Patent No. 6057430
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: No. 6057430el process for the production of
; biologically active dimeric protein
; NUMBER OF SEQUENCES: 14
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/123,233
; FILING DATE:
; CLASSIFICATION: 530
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 339 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: E. coli LC137/pPLm.htGF-beta1 (DSM 5656)
; FEATURE:
```


MOLECULE TYPE: cdna

US-09-169-768-9

Query Match 22.7%; Score 308.6; DB 4; Length 3541;
Best Local Similarity 94.4%; Pred. No. 1.1e-58;
Matches 320; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 896 GCCCTGGATACCACTACTGCTTCCAGCTCCACGGAGAGAACTGCTCGTGGCGGACGCTC 955
DB 3197 GCCCTGGACACCACTATTGCTTCCAGCTCCACGGAGAGAACTGCTCGTGGCGGACGCTG 3256
QY 956 TACATTGACTTCCCGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAAGGGCTACCAT 1015
DB 3257 TACATTGACTTCCCGAAGGACCTGGGCTGGAAGTGGATTTCACAGACCCAAAGGGCTACCAT 3316
QY 1016 GCCAATTTCTGCCCTGGGGCCCTGTCCCTACATCTGGAGCCCTAGACACTCAGTACAGCAAG 1075
DB 3317 GCCAATTTCTGCCCTGGGGCCCTGTCCCTACATTTGGAGCCCTGGACACGCACTACAGCAAG 3376
QY 1076 GTCCTGGGCTGTACACAGCAGCAACCCGGGCGCTCGGGGCGCGCTGCTGGTGGCGG 1135
DB 3377 GTCCTGGGCTGTACACAGCAGCAACCCGGGCGCTCGGGGCGCGCTGCTGGTGGCGG 3436
QY 1136 CAGGCGCTGGAGCCACTGCCCTATGCTGTACTACGTGGGCGCCAAAGCCCAAGGTGGAGCAG 1195
DB 3437 CAGGCGCTGGAGCCACTGCCCTATGCTGTACTACGTGGGCGCCAAAGCCCAAGGTGGAGCAG 3496
QY 1196 CTGTCCAAACATGATGCTGGCTTCCTGCAAGTGCAGCTGA 1234
DB 3497 CTGTCCAAACATGATGCTGGCTTCCTGCAAGTGCAGCTGA 3535

RESULT 15

US-08-868-452-29
Sequence 29, Application US/08868452C
Patent No. 6352972

GENERAL INFORMATION:

APPLICANT: Marcel E. Nimni
APPLICANT: Frederick L. Hall
APPLICANT: Lingtao Wu
APPLICANT: Bo Han Shors
TITLE OF INVENTION: BONE MORPHOGENETIC PROTEINS AND THEIR
TITLE OF INVENTION: USE IN BONE GROWTH
FILE REFERENCE: 17972-11
CURRENT APPLICATION NUMBER: US/08/868,452C
CURRENT FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 29
LENGTH: 337
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(336)

US-08-868-452-29

Query Match 22.6%; Score 308.2; DB 4; Length 337;
Best Local Similarity 94.7%; Pred. No. 8.4e-59;
Matches 319; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 896 GCCCTGGATACCACTACTGCTTCCAGCTCCACGGAGAGAACTGCTGGTGGCGGACGCTC 955
DB 1 GCCCTGGACACCACTATTGCTTCCAGCTCCACGGAGAGAACTGCTGGTGGCGGACGCTG 60
QY 956 TACATTGACTTCCCGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAAGGGCTACCAT 1015
DB 61 TACATTGACTTCCCGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAAGGGCTACCAT 120
QY 1016 GCCAATTTCTGCCCTGGGGCCCTGTCCCTACATCTGGAGCCCTAGACACTCAGTACAGCAAG 1075
DB 121 GCCAATTTCTGCCCTGGGGCCCTGTCCCTACATTTGGAGCCCTGGACACGCACTACAGCAAG 180

QY 1076 GTCTTGGCTCTGTACAAACAGCACAACCCGGGCGGCTCGGGCGCGCCCTGTGCTGCTGCCG 1135
DB 181 GTCTTGGCCCTGTACAAACAGCATAAACCCGGGCGGCTCGGGCGCGCCCTGTGCTGCTGCCG 240
QY 1136 CAGGCGCTGGAGCCACTGCCCATCGTGTACTAGTGGGCGCCGAAGCCCAAGGTGGAGCAG 1195
DB 241 CAGGCGCTGGAGCCACTGCCCATCGTGTACTAGTGGGCGCCGAAGCCCAAGGTGGAGCAG 300
QY 1196 CTGTCCAAACATGATGCTGGCTTCCTGCAAGTGCAGCT 1232
DB 301 CTGTCCAAACATGATGCTGGCTTCCTGCAAGTGCAGCT 337

Search completed: October 9, 2003, 18:35:53
Job time : 111 secs

Db 957 AGCTGGTGAAGCGGAAGCG-ATCGAGGCCATCGCGGCCAGATCTGTCTCAAGCTGCGGC 1016
 QY 186 TTCCAGCCCCCGAGCCCA3GGGAGCTGCGCGCCCGCGCTGCTGAGGAGTACTGG 245
 Db 1017 TCCCAGCCCCCGAGCCCA3GGGAGGTTGCGCGCCCGCGCTGCGCGAGGCGTGTCTG 1076
 QY 246 CTCTTTTACAACTACCGG5GACCGGCTAGCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
 Db 1077 CCTGTACAAACAGCCCG5GACCGGCTGCGCGGGAGAGTGTGAGAACCGGAGCCCGAGC 1136
 QY 306 CAGAGCGGACTACTACGG5AAGGAGTCAACCGCGTGTCTAATGTGGAAGCGGCAACC 365
 Db 1137 CTGAGCGGACTACTACGG5AAGGAGTCAACCGCGTGTCTAATGTGGAAGCGGCAACC 1196
 QY 366 AATCTATGATAAATTCAG3GGGACCCCGCCACAGCTTATATATGCTGTTCACACAGCTGG 425
 Db 1197 AATCTATGATAAATTCAG3GGGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256
 QY 426 AGCTCCGGGAAGCGGTGCG3GAACTGTATTGCTCTCTCGGCGAGAGCTGCGCTGCTGA 485
 Db 1257 AGCTCCGAGAGCGGTACCTGGAACCGGTGTGCTCTCCCGGCGAGAGCTGCGTCTGTA 1316
 QY 486 GGCTCAAGTTAAAGTGA3CAGCAGCTGGAGGTATACCAAGAAATACAGCAATGATTCCT 545
 Db 1317 GGCTCAAGTTAAAGTGA3CAGCAGCTGGAGGTATACCAAGAAATACAGCAATTCCT 1376
 QY 546 GCGCTACCTCAGCAACCGG5CTGTCGCCCCAGTGACTCACCGGAGTGGCTGCTCTTGG 605
 Db 1377 GCGATACCTTCAGCAACCG5CTGTCGCCCCAGTGACTCGCGAGAGTGTATCTCTTTG 1436
 QY 606 ATGTCACCGGAGTGTGCG5CAGTGGCTGACCCCGCAGAGAGGCTATAGAGGTTTTTCGCC 665
 Db 1437 ATGTCACCGGAGTGTGCG5CAGTGGTGTAGCGGTGAGGGGAAATAGAGGCTTTTCGCC 1496
 QY 666 TCAGTGCCCACTCTTCTCTGACAGCAAGATACACACTCCAGCTGGGAATTAACGGGT 725
 Db 1497 TTAGCGCCCACTGCTCTGTGACAGAGGATACACACTGCAAGTGGAGATCAACGGGT 1556
 QY 726 TCAATTCTGCGCGCGGTGAGCTGCGCCACCAATTCAGGCAATGAACCGGCCCTTCTCTG 785
 Db 1557 TCAGTACCGCGCGGAGTGTGCTGCGCCACCAATTCAGGCAATGAACCGGCCCTTCTCTG 1616
 QY 786 TCCTCATGCGCAACCCCGCTGAGAGGCGCCAGACCTGTCAGCTCCCGCAGCCCGCGAG 845
 Db 1617 TTCTCATGCGCAACCCCGCTGAGAGGCGCCAGCATCTGCAAGCTCCCGCAGCCCGCA- 1675
 QY 846 CCCTGGATACCAACAGCTACCCATACGAGTGCAGCTACGCTATCTGTGGCCCTGGATA 905
 Db 1676 -----GCCCTGGACA 1685
 QY 906 CCAACTACTGCTTTCAGCTCAACGAGAGAACTGCTGCGTGCAGGAGCTCTACATTGACT 965
 Db 1686 CCAACTATTGCTTTCAGCTCAACGAGAGAACTGCTGCGTGCAGGAGCTGTACATTGACT 1745
 QY 966 TCGGGAAGGACCTGGGTGTAAGTGGATTATGAACCCCAAGGCTACCATGCCAATTTCT 1025
 Db 1746 TCGGCAAGGACCTGGGTGTAAGTGGATTATGAACCCCAAGGCTACCATGCCAATTTCT 1805
 QY 1026 GCTTGGGCGCTGTGCTTACATCTGAGGCTAGACACTCAGTACAGCAAGTCTCTGGCTC 1085
 Db 1806 GCTTGGGCGCTGTGCTTACATCTGAGGCTAGACACTCAGTACAGCAAGTCTCTGGCTC 1865
 QY 1086 TGTACAAACGACAAACCCGCGCGCTGCGCGCGCGCTGCTGCTGCGCGAGCGCTGG 1145
 Db 1866 TGTACAAACGACATAACCCGCGCGCTGCGCGCGCGCTGCTGCTGCGCGAGCGCTGG 1925
 QY 1146 AGCCACTGCCCATGCTGTACTACGTGGCGCGCAAGGCCCAAGGTGGAGAGCTGTCCAACA 1205
 Db 1926 AGCCGCTGCCCATGCTGTACTACGTGGCGCGCAAGGCCCAAGGTGGAGAGCTGTCCAACA 1985
 QY 1206 TGATGCTGCTTCTGCAATGAGCTGAGCGCGCGCGCTGCTGCTGCGCGAGCGCTGG 1265
 Db 1986 TGATGCTGCTTCTGCAATGAGCTGAGCGCGCGCGCTGCTGCTGCGCGAGCGCTGG 2045

QY 1256 GCGAGCGCGCGCCACCCCGCGCGCT-----CACCGGGCTGTATTTAAGGACA 1317
 Db 2046 GCGCGCGCGCCACCCCGCGCGCGCTGCTTGGCCATGGGGCTGTATTTAAGGACA 2105
 QY 1318 -TCGTGCCCGCCAGCCCACTTGGATCGAATAAA 1349
 Db 2106 CCGTGGCGCCAGCCCACTGGGCGCCATTA 2138
 RESULT 3
 US-10-087-268-4
 ; Sequence 4, Application US/10087268
 ; Publication No. US20030119010A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jonsson, Julie Ruth
 ; APPLICANT: Powell, Elizabeth Ellen
 ; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or condit
 ; FILE REFERENCE: Fibrosis
 ; CURRENT APPLICATION NUMBER: US/10/087,268
 ; CURRENT FILING DATE: 2002-03-01
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 1821
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: {1}..(511)
 ; OTHER INFORMATION:
 ; NAME/KEY: CDS
 ; LOCATION: (512)..(1684)
 ; OTHER INFORMATION:
 ; NAME/KEY: sig_peptide
 ; LOCATION: (512)..(598)
 ; OTHER INFORMATION:
 ; NAME/KEY: 3'UTR
 ; LOCATION: (1685)..(1821)
 ; OTHER INFORMATION:
 US-10-087-268-4
 Query Match 71.1%; Score 968.2; DB 14; Length 1821;
 Best Local Similarity 84.9%; Pred. No. 8,1e-251;
 Matches 1133; Conservative 0; Mismatches 143; Indels 59; Gaps 2;
 QY 6 CCGAGATGCGGCTTCGGGGCTGCGGCTGCTTGGCCGCTGCTGCGCGCTGCTGGCTGC 65
 Db 507 CCCCCATGCGGCTTCGGGCTGCGGCTGCTGCGGCTGCTGCTACCGCTGTGTGGCTAC 566
 QY 66 TAGTGCTGACGCTGCGCGCGCGCGGCGGAGCTGTCCACCTGCAAGACCATCGACATGG 125
 Db 567 TGGTGTGACGCTGCGCGCGCGCGGCGGAGTATCCACCTGCAAGACTATCGACATGG 626
 QY 126 AGCTGTCAAGCGGAGCGCATCGAGGCGCATTCGCGGCGAGATTCTGTCCAAAGCTTCGGC 185
 Db 627 AGCTGTGAAGCGGAGCGCATCGAGGCGCATTCGCGGCGAGATCTGCTGCTTCCAAAGCTGCGGC 686
 QY 186 TTGCCAGCGCGCGGAGCGAGGAGCTGCGCGCGCGCGCGCTGCTGCTGAGGCGACTACTGG 245
 Db 687 TCGCCAGCGCGCGGAGCGAGGAGTGTGCGCGCGCGCGCGCTGCTGCTGAGGCGCGCTGCTG 746
 QY 246 CTCTTTTACAACTACCGCGGAGCGGCTAGCGGGGAAAGTGTGCAACCGGAGCGCGAGC 305
 Db 747 CCTGTACAAACAGCCCGGAGCGGCTGCGCGGAGAGTGTGCAAGCGGAGCGCGAGC 806
 QY 306 CAGAGCGGAGTACTACGCAAGGAGGTGACCCGCGTGTCTAATGTGGAAGCGGCAACC 365
 Db 807 CTGAGCGGAGTACTACGCAAGGAGGTGACCCGCGTGTCTAATGTGGAAGCGGCAACC 866
 QY 366 AATCTATGATAAATTCAGGAGGCAACCCCGCGCTGTATATGCTGTTCACACGCTGG 425
 Db 867 AATCTATGATAAATTCAGGAGGAGTGTACACAGCATATATATGCTTCTTCAACATCAG 926

QY 790 CATGCCCAACCCCGCTGGAGAAGGCCCCAGCACCTGCACAGCT 830
|||||
Db 1302 CATGCCCAACCCCGCTGGAGAAGGCCCCAGCATCTGCAAGCT 1342

```

RESULT 9
US-09-911-904-167
; Sequence 167, Application US/09911904
; Publication No. US2003009623A1
; GENERAL INFORMATION:
; APPLICANT: Fart, Spencer B.
; APPLICANT: Pickett, Gavin G.
; APPLICANT: Neft, Robin Eileen
; APPLICANT: Dunn, II, Robert Thomas
; TITLE OF INVENTION: CANINE TOXICITY GENES
; FILE REFERENCE: 400742000200
; CURRENT APPLICATION NUMBER: US/09/911.904
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/220,057
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: {1}...(489)
; OTHER INFORMATION: n = A, T, C or G
US-09-911-904-167

```

Query Match	26.0%	Score	354.4	DB	11	Length	489
Best Local Similarity	84.1%	Prod. No.	1.3e-85				
Matches	445	Conservative	0	Mismatches	32	Indels	52
Gaps	2						
QY	774	GGCCCTTCCTGCTCCTCATGCCACCCCGCTGGAGAGGCCCGACAGCTCTGCAGAGTCCC	833				
DB	1	GACCCCTTCCTGCTCCTCATGCCACCCACCTGGAGAGGCCCGACAGCTCTGCAGAGTCCC	60				
QY	834	GGCACCCCGGAGGCCCTGGATACCAACAGCTACCCATACGAGTGTCCAGACTACGCAATCTC	893				
DB	61	GGCAGCGCGC-----	70				
QY	894	TGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAAGTCTGCTGGCGCAGC	953				
DB	71	-GGCCCTGGACAACTACTGCTTCAGCTCCACGGAGAAAGTCTGCTGGCTGGCGCAGC	129				
QY	954	TCTACATTGACTTCGCGAAGCACCTGGCTGGAAAGTGATTCTATGAACCCAAAGGCTTACC	1013				
DB	130	TCTACATTGACTTCGCGAAGCATCTGGCTGGAAAGTGATCCATGAGCCCAAGGCTTACC	189				
QY	1014	ATGCAATTTCTGCTGGGCGCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCA	1073				
DB	190	ACGCTAATTTCTGCTGGGCGCTGTCCCTACATTTGGAGCCTGGACACGCACTACAGCA	249				
QY	1074	AGTCTCTGGCTCTGTACAAACAGCACAAACCGGCGCGTGGCGGCGCGCTGCTGCCTGC	1133				
DB	250	AGTCTCTGGCCCTGTACAAACAGCACAAACCGGCGCGTGGCGGCGCGCTGCTGCCTGC	309				
QY	1134	CGCAGGCGCTGGAGCCTGCGCCATCGTGTAAGTCTGCTGGGCGCGCAAGCCAAAGTGGAGC	1193				
DB	310	CGCAGGCGCTGGAGCCTGCGCCATCGTGTAAGTCTGCTGGGCGCGCAAGCCAAAGTGGAGC	369				
QY	1194	AGCTGTCCAACATGATCGTGCATTCTCTCAAGTACAGTGAAGGCCCGGCCCG-CCCACA	1252				
DB	370	AGCTGTCCAACATGATCGTGCCTCTGCAAGTACAGTGAAGGCCCGGCCCGTCCGCA	429				
QY	1253	GGCCCGCCCAACCGGAGGCCCGGGCCCAACCCCGCGCGCTCACCGGG	1301				
DB	430	GGCCCGCCCAACCGGAGGNCGGCGCGCGCGCGCGCTGCTGCCTGGG	478				

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

[illegible]

QY	1136	CAGGCGCTGGAGCCACTGCCATCTGTGTACTACGTGGGCGCAGCAAGTGGAGCAG	1199
Db	241	CAGGCGTGGAGCGGCTGCATCTGTGTACTACGTGGGCGCAGCAAGTGGAGCAG	300
QY	1196	CTGTCCAACATGATCGTGCTTTCTTCTGCAAGTCGAGCTGA	1234
Db	301	CTGTCCAACATGATCGTGCTCTCTGCAAGTCGAGCTGA	339
 RESULT 11 US-09-906-158-3 ; Sequence 3, Application US/09906158 ; Publication No. US20030078217A1 ; GENERAL INFORMATION: ; APPLICANT: Brett P. Monia ; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR-BETA ; FILE REFERENCE: RPS-0257 ; CURRENT APPLICATION NUMBER: US/09/906,158 ; CURRENT FILING DATE: 2001-07-14 ; NUMBER OF SEQ ID NOS: 168 ; SEQ ID NO 3 ; LENGTH: 2574 ; TYPE: DNA ; ORGANISM: Homo sapiens ; FEATURE: ; NAME/KEY: CDS ; LOCATION: (254)...(1492) US-09-906-158-3			
Query Match 18.1%; Score 245.8; DB 11; Length 2574; Best Local Similarity 53.3%; Pred. No. 3.3e-56; Matches 666; Conservative 3; Mismatches 532; Indels 51; Gaps 5;			
QY	36	TGCGCGTGTGCTGTCGGCTCTGTGGCTGTAGTGTGAGCCCTGGCGCGCCGCCCGC	95
Db	261	TGCACCTTGCAAAGGGCTCTGTGGTCTCTGSCCTGCTGAACTTTGCCACGCTCAGCCTCT	320
QY	96	GACTGTCCACTCCGACAGACCATGCACATGAGCTGGTGAAGCGGAAGCGCATGAGGCCA	155
Db	321	CTCTGTGCCACTTGCCACCACCTTGAGCTTGCGGCACATCAAGAAGAAGAGGTTGGAAGCCA	380
QY	156	TTCCGCGCCAGATTCTGTCTCAAGCTTTCGGCTTGCCAGCCCCCGCAGCCAGGGGACGTGC	215
Db	381	TTAGGGGACAGATCTTGAGTAAGCTCAGGCTCACCGACCCCTTGACCCAACGSGTGATGA	440
QY	216	CGCGCGCCCGCTGCGCTGAGCGTAGTACTTGCTCTTTACAACAGTACCCGCGACCGGGTAG	275
Db	441	CCCACG-----TCCCGCTATCAGTCTCTGCGCCCTTTACAACAGCACCCGGAGCTGCTGG	494
QY	276	CCGGGGAAAATGTCGACCGAGCCCCG-----AGCCAGAGGCGGCACTACT	320
Db	495	AGGAGATGCATGGGGAGAGGAGGAAGGCTGACCCAGAAAAACACCGATCGGAATACT	554
QY	321	ACGCCAAGGAGGTACCCCGTGTGTATGTGGAAAGCGGCAACCAATCTATGATAAAT	380
Db	555	ATGCCAAGAATACCATAAATTCACATGATCCAGGGGCTGGCGGAGCACAACGAATGG	614
QY	381	TCAAGGCGACCCCGCACAGTTTTATATGTGTTCACACAGTCGGAGCTCCGGGAACGG	440
Db	615	CTGTCTGCCCTTAAAGGAATTAACCTCCCAAGGTTTTCCCGCTTCAATGTCTCTCAGTGGAGA	674
QY	441	TGCGGGAACCTGTATTCTCTCTCGGCGAGAGCTGCGCCTTCTAGAGCTCAAGTTAAAG	500
Db	675	AAATAGAACCAACCTATT--CCGAGCAGAAATTCGCGGGTCTTTCGCGGTGCCAACCCAGCT	734
QY	501	TGGAGCAGCACGTGGAG-----CTATACCAAAATACAGCAATG	539
Db	735	CTAAGCGGATGACACAGATCGAGCTCTCCAGATCTTCGCGCAGATGAGGCACATG	794
QY	540	ATTCTCTCGGCTACCTCAGTAACCGGCTGCTGGCGCCCGCAGTGACTACCCGAGTGGCTGT	599

Query Match	17.6%	Score	239.6;	DB	12:	Length	4382;		
Best Local Similarity	52.9%;	Pred. No.	1.7e-54;						
Matches	661;	Conservative	0;	Mismatches	544;	Indels	45;	Gaps	5;
<hr/>									
QY	26	CTGCGGCTCTTGGCGCTGCTGCTCGCGCTGCTGCTAGTGTGCTGCTGACGCGCTGGCGG	85						
Db	455	CTGCAGCCATGCACTTGCAGAGGGCTCTGGTAGTCTGTGGCCCTGCTGAATTTGGCCACA	514						
QY	86	CGCGCGCGCGGACGTGTCACCTGTCAAGACCATCGACATGGAGCTGCTGAAGCGGAAGCGC	145						
Db	515	ATCAGCCTCTCTGTCTCCTTTGCACCACGCTTGACCTTCGGCCACATCAAGAAGAAGG	574						
QY	146	ATCGAGGCCATTGCGGCCAGATTCTGTCCAAAGCTTCGGCTTTCGCAGCCCCCGGACGCG	205						
Db	575	GTGGAAAGCAATTAGGGGACAGATTCTGAGCAAGCTCAGGCTCACCAGCCCCCTTGAGCCA	634						
QY	206	GGGGACGTGCGCGCGCGCGCGCTGCTTGGGCACTACTGGCTCTTTTAAACAGTACCGCG	265						
Db	635	TCGGTGTGACCCACG-----TCCCTATCAGGTCTCTGGCATTTCACACAGCACCGCG	688						
QY	266	GACCGGCTAGCGC-----GGGAAAGTGTCTAAACCGGAGCGCGGACCGAG	310						
Db	689	CAGTTGCTGGAAGAGATGCACGGGACAGCGAAGGCTGCACCTCAGAGACCTTCGCA	748						
QY	311	GGCGACTACTACGGCAGGAGGTACCCCGCGTCTTAATGGTGAAGAGCGGCACCAATC	370						
Db	749	CTGAGTACTATGCCAAGAGATCCATAATTCACATGATCAGGGACTTCGGCGAGCAC	808						

Qy	371	TATGATAAATTC	AAAGGGCA	CCCCC	CACAGCTT	ATATATGCTTT	CAACAGTCGGAGCTC	430	
Db	809	AATGAAGCT	CGCCGTCTGCC	CAAAAGGAAT	TACCTCT	TAAGGTTTTT	TGTTTCAATGTGTCC	868	
Qy	431	CGGAAGCGGT	CGCGNAAC	TGTATTGCTCT	CTCGGSCAGAGT	CGCCTGTGCTGAGGCTC	490		
Db	869	TCAGTGAGAAAAT	GGAA	CAATCTGTT	CCGGGCAGAGTT	CGGGTCTTT	CGGGTGTGCC	928	
Qy	491	A	-----	AGTTAA	AGTGGAGCAGCAGCT	GGAGCTATACCA	AGAAA	529	
Db	929	AAOCCAGCT	CCAAGCGCA	AGAGCAGAGAAT	TGAGCTCTT	CCAGATACTT	CGACCGGAT	988	
Qy	530	TACAGCAAT	GATTCCTGG	CGCTPACCT	CAGCAACGGCT	GCTGCCCCC	CAGTGA	589	
Db	989	GAGCACA	TACCCAAAGCAG	CGCTCAT	AGTGTGGCAAGAAAT	CTGCCCA	AAAGGGCACC	1048	
Qy	590	GAGTGGCT	GCTCTTGTAT	GACCGGAGTTGT	GCGGCAGTGGCT	GACCCG	CAGAGAGGCT	649	
Db	1049	GAATGGCT	GCTCTTCGAT	GTCACTGACACT	GTGCGCAGTGGCT	TTTGAGG	AGAGAGTCC	1108	
Qy	650	ATAGAGGG	TTTTCGCCTCA	TGCCCCACT	TTCTCTG	CAGCAAGAA	ATATACACACT	709	
Db	1109	AACTTGG	GTCTGGAANAATCA	SCATCCCACT	GTCCATGTCA	CACACCTTT	CAGCCCAATGGAGC	1168	
Qy	710	GTGCAAA	TAAACG	GGTGTAA	TTCTGCCCCCGGGTGACCT	GCCACCA	TTTCCAGGCT	768	
Db	1169	ATACTG	AAATATGTT	CATGAGGT	GATGGAAATCA	MAATTC	AAAGGATGGACAT	1228	
Qy	769	GAACCG	GCCCTTCTGCTC	TTCTAT	TGGCCACCCCGCTGG	AGAGGGCC	CAGCACTGCACAG	828	
Db	1229	GACCAT	GGCGTGTGAGACCT	GTGGGCGGTCT	CAAGAAGCA	AAAGGATC	CACCAACACCCACAC	1288	
Qy	829	CTCCG	CACCGCGCAG	CCCTGGAT	ACCAACAGCTAC	CCCATAC	GAGTGCAGCACTACGC	888	
Db	1289	CTGAT	CTCATGATGATCT	CCCCAC	CCGACTGGACAGCC	CAGGCGCAT	GCAGAGG	1348	
Qy	889	ATCTCT	--	GGCCCT	GGATACCA	ACTACTGCTT	CAGCTCC	CAGGAGAGA	946
Db	1349	AAGAAG	AGGGCCCTGGAC	ACAATTA	GCTTCTCCGCA	AACTGGAGG	AGAACTGCTGTGTA	1408	
Qy	947	CGGCAG	CTCATTTGACT	TTCCGAGAG	CACTGGGCT	GGAAGTGA	TTTCAT	GAACCCAAG	1006
Db	1409	CGCCCC	CTTATATGTGACT	CCGGCAG	GATGCTAGGCT	GTGAAATGGG	TCCACAGAACTTAAG	1468	
Qy	1007	GGCTAC	CATGCAATTTCTG	CTGGGCC	CTGTCCCT	TACATCT	GTGAGCCT	TAGACACT	1066
Db	1469	GTTTACT	ATGCCAACTCT	GCTCAGG	CCCTTGCC	CATACCT	CGCAGCGCAG	ACAACAAC	1528
Qy	1067	TACAG	AAGTTCCTGGCTC	GTAC	ACAACAGCA	CAAAACCC	GGGCGGT	CGCGGGCGCGTGC	1126
Db	1529	CATAG	CAGGTGTCTGGAC	TTATAC	AACCCCTGA	ACCAGAG	CGCTGTGCT	GTGCCATGC	1588
Qy	1127	TGCT	GTCCGAGGCGCT	GTGAGCC	ACTGCC	CATCTGT	ACTACGT	GCGGCCACAG	1186
Db	1589	TGCGT	CCCCCAGGACCT	GTGAG	CCCTTAC	CACTTTGT	TACTAT	GTGGCAGAA	1648
Qy	1187	GTGAG	CAGCT						

RESULT 14
US-09-906-158-10
; Sequence 10, Application US/09906158
; Publication No. US20030078217A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR-BETA 3 EXPRESSION
; FILE REFERENCE: RTS-0257
; CURRENT APPLICATION NUMBER: US/09/906,158
; CURRENT FILING DATE: 2001-07--4
; NUMBER OF SEQ ID NOS: 168

```

: SEQ_ID NO 10
: LENGTH: 2879
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (611)...(1843)
JS-09-906-158-10

Query Match      17.4%  Score 236.4;  DB 11:  Length 2879;
Best Local Similarity 52.8%  Pred. No. 1.1e-53;
Matches 656;  Conservative 0;  Mismatches 541;  Indels 45;  Gaps 5;

QY      36  TGGCGTCTGCTGCGCGTCTGCTGGCTGCTAGTGTCTAGCGCCTGCGCGCGCGCGCGCG 95
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      612  TGCACTTGCANAAGGCTCTGGTAGTCTGCGCCTGCTGAACCTTGGCCACAATCAGCCTCT 671
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      96  GACTGTCCACTGCCAAGACCATCCACATGGAAGCTGTGTGAAGCGGACGCGATCGAGGCCA 155
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      672  CTCTGTCCACTTGCACCACTTGGAGCTTGGCCACACATCAAGAAGAAGAGGGTGGAAAGCCA 731
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      156  TTGCGGCCAGATTCTCCAAAGCTTTCGGCTTCCAGCCCCCGGAGCCCGGAGGCGAGCTGC 215
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      732  TTAGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCATCGGTGATGA 791
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      216  CGCGCGCCCGCTGCGCTGAGGCAAGTACTGCTCTTTACAAAGTACCGGAGACCGGCTAG 275
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      792  CCCAGC-----TCCCTATCAGTCTCTGCACTTTACAAAGACACGCGGAGTTGCTGG 845
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      276  CCG-----GGGAAGTCTCGAAACCGGAGCCCGAGCCAGGACGAGCGGAGCTACT 320
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      846  AAGAGATGCAAGGGGAGGAGGAAGGCTGCACTCAGGAGACTCGGAGTCTTGAGTACT 905
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      321  ACGCCAAGGAGTCAACCGGCTGCTAATGTGTGAAGGCGCAACCAAACTATGATAAAT 380
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      906  ATGCCAAGAGATCCATAAAATTGCACATGATCCAGGACTCGGCGGAGCACAATGAATCG 965
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      381  TCAAGGCAACCCCAACAGCTTATATGCTGTTCACACGTCGGAGCTCCGGGAAGCGG 440
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      966  CCGTCTGCCCAAGGAATTAACCTCTAAGGTTTTCTGTTCAATGTCTCTAGTGGAGA 1025
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      441  TGCGGGAACCTGTATGCTCTCTCGGCGAGCTGCGCTGCTGTGAGCTCA----- 491
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1026  AAAATGAAACCAATCTGTCTCGGCGACAGTTCCGGGTCTTCGGGTGCCCAACCCAGCT 1085
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      492  -----AGTTAAAGTCGAGCAGCAGCTGGAGCTATACCAGAAATACAGCAATG 539
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1086  CCAAGCGGACAGCAGCAGAATTGAGCTCTTCCAGATACTTCGACCGGATGAGCACATAG 1145
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      540  ATTCTTGGCGTACCTCAGCAACCGGTGCTGGCGCCCACTGACTCACCAGAGTGCTGT 599
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1146  CCAAGCAGCGTACATAGTGGCAAGAATCTGCCCAAGGGCACCGCTGAATGGCTGT 1205
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      600  CCIITGATGCACGGAGTTGTGGCGCAGTGGCTGACCCGACAGAGGCTATAGAGGTT 659
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1206  CTTTCGATGTCACAGACTGTGCGCGAGTGGCTGTGAGGAGAGAGTCCAACTTGGGTC 1265
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      660  TTCGCCTCAGTGGCCACTCTTCCTCTGACAGCAAGATAACACACTCCAGCTGGAATTA 719
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1266  TCGAAATCAGCATCCACTGTCCAATGTCACACTTTTCAGCCCAATGGAGACATACTGAAA 1325
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      720  ACG--GGTTCAAATTCGCGCGCGGGGTGACCTGGCCACCAATCAGGCGATGAACCGGCC 778
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1326  ATGTTCAATGAGGTGATGGAAATCAAAATTCAAAGAGGTGGACAATGAAGATGACCAATGCC 1385
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      779  TTCCTGTCTCTCATGGCCACCCCGCTGGAGAGGCCCAAGCACTTGCACAGCTCCCGGCAC 838
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1386  GTGAGACCTTGGGGCTCTCAAGAAACAAAGGATCACCAACACCCACACCTGATCTCTCA 1445
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      839  CGCGAGCCCTGGATACCAACAGCTACCCATACGACGTGCGAGACTACGACTCTCT--GG 896
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1446  TGATGATCCCCCAACACCGATCTGCACAGCCCCAGGCCAGGCGCTCAGCAGGAAGAAGAGG 1505
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2003, 16:00:18 ; Search time 3216 Seconds
(without alignments)
10285.570 Million cell updates/sec

Title: US-10-017-372E-36
Perfect score: 1361
Sequence: 1 tggatccgagatggcgcctt.....cgattaaagcgcgcgact 1361

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12:52238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estnu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_hic:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_hic:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrt:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	703.6	51.7	1072	12	BM562135
c 2	698.4	51.3	1201	13	BM562135 AGENCOURT
3	688.8	50.6	900	13	BM562135 AGENCOURT
c 4	687.4	50.5	983	13	BM562135 AGENCOURT

c 5	646	47.5	1041	13	BM562135
c 6	607.2	44.6	888	13	BM562135
c 7	607	44.6	1043	12	BM562135
c 8	600	44.1	950	13	BM562135
c 9	577.4	42.4	859	9	AL530081
c 10	570.6	41.9	713	10	BE312000
c 11	562.6	41.3	717	10	BE260971
c 12	562	41.3	902	13	BM562135
c 13	534.4	39.3	925	12	BM562135
c 14	521.8	38.3	841	9	AL530081
c 15	521.4	38.3	956	12	BM562135
c 16	513.8	37.8	871	13	BM562135
c 17	506.6	37.2	1013	12	BM562135
c 18	506.6	37.2	1093	12	BM562135
c 19	504.8	37.1	943	12	BM562135
c 20	502	36.9	821	10	BM562135
c 21	501.8	36.9	773	12	BM562135
c 22	498.8	36.6	1049	12	BM562135
c 23	493.2	36.2	757	14	CA309731
c 24	486	35.7	1047	13	BM562135
c 25	472.8	34.7	675	10	BE261784
c 26	466.2	34.3	727	9	AL131171
c 27	464.2	34.1	960	13	BM562135
c 28	461	33.9	785	9	AL148173
c 29	459.8	33.8	637	12	BM562135
c 30	457.2	33.6	565	14	CD287835
c 31	455.8	33.5	748	10	BM562135
c 32	455.4	33.5	814	12	BM562135
c 33	452.2	33.2	548	12	BM562135
c 34	450.6	33.1	722	13	BM562135
c 35	450	33.1	982	13	BM562135
c 36	448.8	33.0	928	13	BM562135
c 37	448.6	33.0	695	14	CA425775
c 38	447.6	32.9	981	12	BM562135
c 39	446.6	32.8	697	13	BM562135
c 40	446	32.8	659	10	BM562135
c 41	446	32.8	700	13	BM562135
c 42	445.4	32.7	690	13	BM562135
c 43	445.4	32.7	697	14	CA426391
c 44	444.6	32.7	778	9	AL148173
c 45	443.6	32.6	1015	10	BM562135

ALIGNMENTS

RESULT 1
BM562135
LOCUS
DEFINITION
AGENCOURT_6562032 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745463
5', mRNA sequence.
ACCESSION
BM562135
VERSION
BM562135.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 1072)
AUTHORS
NIH-MGC
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNL12768 row: h column: 09
High quality sequence start: 9

1072 bp
linear
EST 20-FEB-2002
BM562135
NIH-MGC
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNL12768 row: h column: 09
High quality sequence start: 9

High quality sequence stop: 647.

FEATURES		Location/Qualifiers	
source		1. .1072	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5745463"	
		/tissue_type="leukocyte"	
		/lab_host="DH10B"	
		/clone_lib="NIH_MGC_118"	
		/note="vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."	
BASE COUNT		217 a	356 c 113 g 186 t
ORIGIN			
		Query Match 51.7%; Score 703.6; DB 12; Length 1072;	
		Best Local Similarity 81.6%; Pred. No. 2.3e-143;	
		Matches 886; Conservative 0; Mismatches 144; Indels 56; Gaps 4;	
QY	127	GCTGGTGAAGCGGGAAGCGCA	CGAGGGCCATTGCGGGCCAGATCTGTCCAAAGCTTCGGCT 186
Db	20	GCTGGTGAAGCGGGAAGCGCA	CTAGGCCATCCGGGCGCAGATCTGTCCAAAGCTTCGGCT 79
QY	187	TGCCAGCCCCCGGAGCGAGCGAGTGC	CGCCCGCGCTGCTGTCGACGACTTGGC 246
Db	80	CGCCAGCCCCCGGAGCGAGCGAGTGC	CGCCCGCGCTGCTGTCGACGACTTGGC 139
QY	247	TCTTTACAACTAGTACCCCGGACCGGGTAG	CGGGGAAAGTGTGCAACCGGAGCCGAGCC 306
Db	140	CTGTGTACAACTAGTACCCCGGACCGGGTAG	CGGGGAAAGTGTGCAACCGGAGCCGAGCC 199
QY	307	AGAGCGGCACTACTACGCCAAGGAGGTAC	CCCGCGTCTAATGTGTGAAGCGGCAACCA 366
Db	200	TGAGGCCCACTACTACGCCAAGGAGGTAC	CCCGCGTCTAATGTGTGAAGCCCAACCA 259
QY	367	AATCTATGATAAATTAAGGTCACCCCGGAC	CGTCTATATGCTGTTCACACGTCGGA 426
Db	260	AATCTATGATAAATTAAGGTCACACGAGT	TACACAGCATATATGTTCTTCAACACATCAGA 319
QY	427	GCTCGGGAAGGGTGGCGGACCTGTATT	CTCTCTCGGGCAGAGTGCCTGCTGAG 486
Db	320	GCTCGGGAAGGGTGGCGGACCTGTATT	CTCTCTCGGGCAGAGTGCCTGCTGAG 379
QY	487	GCTCAAGTTAAAGTGGAGCGACGCTGG	AGCTATACCAAGAAATACAGCAATGTCCTG 546
Db	380	GCTCAAGTTAAAGTGGAGCGACGCTGG	AGCTATACCAAGAAATACAGCAATGTCCTG 439
QY	547	GGGTACTCTAGCAACCGGCTGCTGG	CCCCCAGTGTCTACCGGAGTGGCTGCTTGA 606
Db	440	GGATACCTTCAGCAACCGGCTGCTGG	CCCCCAGTGTCTACCGGAGTGGCTGCTTGA 499
QY	607	TGTCACCGAGTGTGGCGGAGTG	GGCTCACCGGAGAGGCTATAGAGGTTTCGCT 666
Db	500	TGTCACCGAGTGTGGCGGAGTG	GGCTCACCGGAGAGGCTATAGAGGTTTCGCT 559
QY	667	CAGTGGCCCACTTCTCTGTGACGACAA	AGATAACACACTCCACGCTGGAAATTAACGGTT 726
Db	560	TAGCGCCCACTTCTCTGTGACGACG	AGATAACACACTCCACGCTGGAAATTAACGGTT 619
QY	727	CAATTCTGGCGCGGGGTGACCTGG	CGCACCACTTACCGGATGACCGGCGCTTCCTGCT 786
Db	620	CACATACCGCGCGGGGTGACCTGG	CGCACCACTTACCGGATGACCGGCGCTTCCTGCT 679
QY	787	CTCATGGCCACCCCGCTGGAGCGG	CCCGACACCTGCACAGCTCCCGGACCGCCGAGC 846
Db	680	TCTCATGGCCACCCCGCTGGAGCGG	CCCGACACCTGCACAGCTCCCGGACCGCCGAGC 739

QY	847	CTTGATACCAACAGCTACCCATACG	AGCTGCGCAGACTAGGCATCTCTGGCCTGGATAC 906
Db	740	-----	-----CCCTGGACAC 749
QY	907	CAACTACTGCTTACGCTCCACG	GAGAGAACTGCTGGTGGGAGCTCTACATTGACTT 966
Db	750	CAACTATTGCTTACGCTCCCGG	AGAAACTGCTGGTGGGAGCTCTACATTGACTT 809
QY	967	CGGGAAGCACTGGCTGGAACTG	GATTTCATGAACCAAGGGCTACCATGCCAATTCTG 1026
Db	810	CGGAGGACACTGGCTGGAACTG	GATTTCATGAACCAAGGGCTACCATGCCAATTCTG 869
QY	1027	CTTGGGGCCCTGTCTTACATCT	TGGAGCTTAGACACTAGTACAGCAAGTCTCTGGCTCT 1086
Db	870	CTTGGGGCCCTGTCTTACATCT	TGGAGCTTAGACACTAGTACAGCAAGTCTCTGGCTCT 929
QY	1087	GTACAACCAAGCACACCGCG	CGCTGCGGGCGCGTGTGCTGCGCGGCGGCTG 1144
Db	930	GTACAACCAAGCACACCGCG	CGCTGCGGGCGCGTGTGCTGCGCGGCGGCTG 989
QY	1145	GAGCCACTGCCCAT--CGT	TACTACGTGGGCGC--AAGCCCAAGGTGAGCAGCTGTC 1200
Db	990	GACCCCTGCCCCCTTCGGG	TACTACGTGGGCGC--AAGCCCAAGGTGAGCAGCTGTC 1049
QY	1201	CAACAT 1206	
Db	1050	CCACAT 1055	

RESULT 2			
LOCUS	BX355682	1201 bp	mRNA linear EST 05-MAY-2003
DEFINITION	BX355682 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1002YJ05 3-PRIME, mRNA sequence.		
ACCESSION	BX355682		
VERSION	BX355682.1	GI:30371987	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0D1002CE03NP1&cluster=9160.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1002CE03NP1.		

FEATURES			
source	1. .1201		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0D1002YJ05"		
	/tissue_type="PLACENTA COT 25-NORMALIZED"		
	/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"		
	/note="1st strand cDNA was primed with a NofI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		
BASE COUNT	198 a	326 c	386 g 253 t
ORIGIN			

[illegible]

Qy	1118	GCGCCGTGTCGCTGGCCGAGCGCTGGACCACTGCCCATCTGTTACTGTTGGGCGCCG	1177
Dd	196	GCGCCGTNGT-CGTGGCG-AGGCGCTGGAGCCGCTGCCATCTGTTACTGTTGGGCTGC	138
Qy	1178	AAGCCAAAGTGTCGAGCAGTGTCACAATGATCGTGGCTTCCTGCAAGTGCACCTCAGGC	1237
Dd	137	AAGCCCAAGTGTCGAGCAGTGTCTCCAATCATGATCGTGGCTTCCTGCAAGTGCACCTCAGGT	78
Qy	1238	CCGCCCCCGCCCACAGCCCGCCACCAGCCGCGAGCCGCGCCACCCCGCCGCGCTCAC	1297
Dd	77	CCGCCCCCGCC-----CCGCCCCCGCCGAGCCGCGCCACCCCGCCGCGCCCGC	23
Qy	1298	CGSGGGTG 1305	
Dd	22	CTGCTTG 15	
RESULT 5			
BX324511/c			
LOCUS BX324511 Homo sapiens: NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens			
DEFINITION cDNA clone CS0C024YD20 3-PRIME, mRNA sequence.			
ACCESSION BX324511			
VERSION BX324511.1 GI:30332:81			
KEYWORDS EST.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
1 (bases 1 to 1041)			
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.			
Full-length cDNA libraries and normalization			
Unpublished			
COMMENT Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 9160.r For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0AC024DB10NP2&cluster=9160.r. Contact :			
Peng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1500			
Paraday Avenue Genoscope sequence ID : CS0AC024DB10NP2.			
Location/Qualifiers			
1..1041			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0D2024YD20"			
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"			
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo(dT)			
primer. Five prime end enriched, double-strand cDNA was			
digested with Not I and cloned into the Not I and EcoR V			
sites of the PCMVSPORT 6 vector. Library was normalized."			
BASE COUNT 176 a 283 c 346 g 219 t 17 others			
ORIGIN			
Query Match 47.5% Score 646; DB 13; Length 1041;			
Best Local Similarity 80.4%; Pred. No. 8.8e-131;			
Matches 868; Conservative 10; Mismatches 135; Indels 67; Gaps 8;			
Qy	279	GGGAAGTGTGAACCGGAGCCCGCAGCGGAGCTACTACGCCAAGGAGGTCAACC	338
Dd	1022	GGGAGAGSAGACCGGAGGCCCGACCTTRASCRC---ACTAYACCCAGAGTAGTACACA	966
Qy	339	GCGTGTAAATGTCGAAAGCGGCAACCAAATCTATGATAATTCAAGGGCACCCCCACA	398
Dd	965	GCGTGYTANTGKT-GAAACCCACCAACGAATCTATGACAGTTCAGCAGAGTACACACA	907

Qy	399	CCTTAATATGCTGTTCTCAACACGCTCGGAGCCTCCGGGAAGCGGTCCCGGAACCTGTTATTGC	458
Db	906	GCATATATATGTTCTTCCACACATACAGCTCCGAGAAAGCGGTACCTGAACCCGCTGTTC	847
Qy	459	TCTCTC-GGGCAGAGCTGCGCCCTGCTGAGGCTCAAGTTAAAGTGGAGCAGCAGCTGGAG	517
Db	846	TCTCCSGGCGAGCTGCGCTGCTGAGGCTCAAGTTAAAGTGGAGCAGCAGCTGGAG	787
Qy	518	CTATACCAGAAATACAGCAATGATTCCTTGGCGCTACCTCAGCAACCGGCTGCTGGCCGCC	577
Db	786	CTGTACCAGAAATACAGCAAAATTCCTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	737
Qy	578	AGTGACTCACCGGAGTGGCTGCTCTTGTATGTCACGGAGTTGTGGCGCAGTGGCTGACC	637
Db	726	AGCGNCTGCCAGAGTGGTTATCTCTTGTATGTCACGGAGTTGTGGCGCAGTGGTTCAGC	667
Qy	638	CGCAGAGAGGCTATAGAGGGTTTTCGCTCAGTGCCCACTCTTCCTCTGACAGCAAGAT	697
Db	666	CGTGGAGGGGAAATGTA-GGCTTTCGCCCTTAGCGGCCACTGCTCTGTGACAGCAGCGAT	608
Qy	698	AACACACTCCAGCTGGGAAATTAACGGGTTCAATTCCTGGCCGCCGGGTGACCTGGCCACC	757
Db	607	AACACACTGCAAGTGGACATCAACGGGTTCACTACCGCGCCGAGGTGACCTGGCCACC	548
Qy	758	ATTCAACGGCATGAACCGGCCCTTCTCTGCTCCTATGGCCACCCGCTCGAGAGGGCCACG	817
Db	547	ATTCAATGCGATGAACCGGCCCTTCTCTGCTCCTATGGCCACCCGCTCGAGAGGGCCACG	488
Qy	818	CACCTGACAGCTCCCGGACGCGGAGCGCTGGATACCAACAGTACCCATCAGACGTG	877
Db	487	CATCTGCAAGCTCCCGGACCGCGGAGCGGCTGGAG-----	452
Qy	878	CCAGACTACGCATCTCTGGCCCTGGATACCAACTACTGCTTCAGCTCCACGAGAGAAGAC	937
Db	451	-----ACCAACTATTGCTTCAGCTCCACGAGCAAGAC	419
Qy	938	TGCTCGTGGCGAGCTCTACATTGACTTCCGGAGGACCTGGCTGGAAGTGGATTCAAT	997
Db	418	TGCTCGTGGCGAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	359
Qy	998	GAACCCAAAGGCTACCATGCCAATTTCTGCTGGGCGCTGTCCCTACATCTCGAGCGTA	1057
Db	358	GAGCCCAAGGCTACCATGCCAATTTCTGCTGGGCGCTGTCCCTACATCTCGAGCGTG	299
Qy	1058	GACACTAGTACAGCAAGTCTCGCTGTGTACACAGCAGCAACCGGGCGCTCGGCG	1117
Db	298	GACAGCGAGTACAGCAAGTCTGGCCCTGTACACAGCAGCATAAACC-GGGCGCTCGGCG	240
Qy	1118	GGCGGCTGCTGCTGCCCGCAGGCGCTGGAGCCACAGCCCATGCTACTAGTGGGCGCG	1177
Db	239	GGCGCGTG-TGCTGTCCGAGGCGCTGGAGCCGCTGCCATCGTACTAGTGGGCGCG	181
Qy	1178	AAGCCCAAGTGGAGCAGCTGTCCACATGATCGTGGCTTCTTGCAGTGCAGCTGAGGC	1237
Db	180	AAGCCCAAGTGGAGCAGCTGTCCACATGATGCTGGCTTCTTGCAGTGCAGCTGAGGT	121
Qy	1238	CCCGCCCGCCACAGCCCGCCACCCCGGAGCGCCGCCACCCCGCCCGCCGCT---	1294
Db	120	CCCGCCCGCCCGCCCGCCCGGAGCGCCGCCACCCCGCCCGCCCGCCCGCTGCC	61
Qy	1295	-----CACCGGGGCTGTATTAAAGGACATCGTGGCCCAAGCCCACTTGGATCGATTAAA	1349
Db	60	TTGCCCATGGGGCTGTATTAAAGGACACCGTCCCAAGGCCACCTGSCACCACTAA	1
RESULT 6			
BX434425/c			
LOCUS			
DEFINITION			
BX434425 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0E014YE16			
3-PRIME, mRNA sequence.			
ACCESSION			
BX434425			
VERSION			
BX434425.1			
KEYWORDS			
EST.			
GI:30779291			
EST.			
888 bp			
mRNA			
linear			
EST 15-MAY-2003			

[illegible]

```
BASE COUNT      202 a      329 c      345 g      154 t      13 others
ORIGIN
Query Match      44.6%; Score 607; DB 12; Length 1043;
Best Local Similarity 80.9%; Pred. No. 2.9e-122;
Matches 734; Conservative 0; Mismatches 122; Indels 51; Gaps 1;
QY 396 ACAGCTTATATATGCTGTTTAAACACGTCGGAGCTCCGGGAAGCGGTGCGGCAACCTGTAT 455
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ACAGCATATATATGTTCTTCAACACATCAGAGCTCGAGAGCGGTACCTGAACCGGTGT 60
QY 456 TGTCTCTCGGGCAGAGCTGCGCTCTGAGGCTCAAGTTAAAGTGGAGCAGCAGTGG 515
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TGTCTCTCGGGCAGAGCTGCGCTCTGAGGCTCAAGTTAAAGTGGAGCAGCAGTGG 120
QY 516 AGCTATACCAGAAATACAGCAATGATCTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCC 575
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 AGCTGTACCAGAAATACAGCAATGATCTCTGGCGTACCTCAGCAACCGGCTGCTGGCAC 180
QY 576 CCAGTGACTACCGGAGTGGTCTCTTTGATGTACCGGAGTGTGCGGCGAGTGGCTGA 635
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CCAGCGACTCGCCAGAGTGTTATCTTTTGTATCTCACCGGAGTGTGCGGCGAGTGTGA 240
QY 636 CCGCGAGAGGCTATAGAGGGTTTTCGGCTCAGTCCGACCTCTCTCTGACAGCAAG 695
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GCGTGGAGGGGAATTTAGAGGCTTTTCGGCTTACCGCCCACTGCTCTGTGACAGAGG 300
QY 696 ATAACACATCCACGTGGAAATTAACGGGTTCAATTTCTGGCGCGCGGGTGTACCTGGCCA 755
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ATAACACATGCAAGTGGACATCAACGGGTTCACTACCGCGCGCGAGGTGACCTGGCCA 360
QY 756 CCATTACGCGATGAACCGGCTCTCTGCTCTGATGGCCACCGCGCTGGAGGGCC 815
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 CCATTACGCGATGAACCGGCTCTCTGCTCTGATGGCCACCGCGCTGGAGGGCC 420
QY 816 AGCACTTGACAGTCCCGGACCGCGGAGCCCTGGATACCAACAGCTACCCATACGAG 875
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 AGCATCTGAAAGTCCCGGACCGCGA ----- 449
QY 876 TGGCAGACTAGGATCTCTGCGCTGTGATACCAACTACTGCTTCAGCTCCAGGAAAGA 935
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 450 -----GCCCTGGACACCAACTATTGCTTCAGCTCCAGCGAAGA 489
QY 936 ACTGCTGCTGCGGACGCTTACATTGACTTCCGGAGGACCTGGGCTGGAAGTGATTC 995
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 ACTGCTGCTGCGGACGCTTACATTGACTTCCGAGGACCTCGGCTGGAAGTGATCC 549
QY 996 ATGAACCCAAAGGCTACCATGCCAATTTCTGCTGGGGCCCTGTCCCTACATCTGAGGCC 1055
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 ACGAGCCAAAGGCTACCATGCCAATTTCTGCTGGGGCCCTGTCCCTACATTTGAGGCC 609
QY 1056 TAGACACTCAGTACAGCAAGTCTGCTGGCTGTACAAACAGCAACACCGGGCGCTCGG 1115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 TGGACACGAGTACAGCAAGTCTGCTGGCTGTACAAACAGCAACACCGGGCGCTCGG 669
QY 1116 CGGCGCGCTGCTGCTGCGCGCAGCGCTGGAGCCACTGCCCATCGTGTACTAGTGGGCC 1175
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 670 CGGCGCGCTGCTGCTGCGCGCAGCGCTGGAGCGGCTGCCCATCGTGTACTAGTGGGCC 729
QY 1176 GCAAGCCCAAGGTGGACAGCTGCTCAACATGATGCTGGCTTCTGCAAGTGACGTGAG 1235
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 730 GCAAGCCCAAGGTGGACAGCTGCTCAACATGATGCTGGCTTCTGCAAGTGACGTGAG 789
QY 1236 GCCCGCGCCCGCCACAGCGCCGCCACCGCGGAGCGCCCGGCCCGCCCGCCCGCTC 1295
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 790 GGGTCCGCGCCCGCCACAGCGCCCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 849
QY 1296 ACCGGGG 1302
   |||||
Db 850 GGGCGCG 856
```

RESULT 8

```
BX383773/c      950 bp      mRNA      linear      EST 08-MAY-2003
LOCUS
DEFINITION      BX383773 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
                  cDNA clone CS0DK001YA15 3-PRIME, mRNA sequence.
ACCESSION      BX383773
VERSION        BX383773.1 GI:30457168
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                1 (bases 1 to 950)
                Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
                Full-length cDNA libraries and normalization
                Unpublished
CONTACT      Genoscope
                Genoscope - Centre National de Sequencage
                BP 191 91006 EVRY cedex - France
                Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
                Library was constructed by Life Technologies, a division of
                Invitrogen. This sequence belongs to sequence cluster 9160.r For
                more information about this cluster, see
                http://www.genoscope.cns.fr/
                cgi-bin/cluster.cgi?seq=CS0DK001AA08NP1&cluster=9160.r. Contact :
                Feng Liang Email: fliang@lifetech.com URL :
                http://fulllength.invitrogen.com/ invitrogen Corporation 1600
                Faraday Avenue Genoscope sequence ID : CS0DK001AA08NP1.
                Location/Qualifiers
                1..950
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DK001YA15"
                /cell_type="HELA CELLS COT 25-NORMALIZED"
                /cell_line="HELA"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                primer. Five prime end enriched, double-strand cDNA was
                digested with Not I and cloned into the Not I and EcoR V
                sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      171 a      260 c      312 g      195 t      12 others
ORIGIN
Query Match      44.1%; Score 600; DB 13; Length 950;
Best Local Similarity 79.2%; Pred. No. 9.5e-121;
Matches 789; Conservative 8; Mismatches 136; Indels 63; Gaps 5;
QY 353 GAAAGCGGCAACCAATCTATGATAAATTCGAAGGTCACGGGACCCACAGCTTATATGCTG 412
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 943 GRAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGMACACACAGCATATATGTTIC 884
QY 413 TTCACACGTCGGAGCTCCGGGAAGCGTCCCGGAACCTGCTGCTCTCGGGCAGAG 472
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 883 TTCACACATAGAGCTCCGAGAACGGGTACCTGACCCGCTGTCTCCCGGCGAGAG 824
QY 473 CTGCGCTGCTGAGGCTCAAGTTAAAGTGGAGCAGCAGCTGGAGCTATACCAAAATAC 532
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 823 CTGCTGCTGCTGAGGCTCAAGTTAAAGTGGAGCAGCAGCTGGAGCTGTACCAAAATAC 764
QY 533 AGCAATGATTCCTGGGCTACCTCAGCAACCGGCTGCTGGCCCGGCTGACTCACCGGAG 592
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 763 AGCAACAAATATCTGGGATACCTCAGCAACCGGCTGCTGGCACCCGAGCTCGCCAGAG 704
QY 593 TGGCTGCTCTTGTATGTACCGGAGTTGTGCGGAGTGGGTGACCCCGCAGAGAGGCTAT 652
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 703 TGGTATCTTTGATGTACCGGAGTTGTGCGGAGTGGTGTAGCCCGTGGAGGGGAATT 644
QY 653 GAGGGTTTTGCGCTCAGTGCCCACTCTTCTCTGACAGCAAGATACACACTCCAGCTG 712
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 643 CA-GGCTTTGCGCTTAGCGCCCACTGCTCTCTGTACAGCAGGGATTACACACTCCAAAGT 585
QY 713 GAAATTAACGGGTTCAATTCTGGCGCGGGGTGACCTGCCACCATTACGGCATGAAC 772
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

Db      584 GACATCAACGGTTTACTACCGCGCGCGAGTGACCTGGCCACCATTATGCGCATGAAC 525
QY      773 CGGCGCTTCTGCTCTTCATCGCACACCCGCTCGAGAGGCGCCAGCACCTGCGACGCTCC 832
Db      524 CGGCGCTTCTGCTCTTCATCGCACACCCGCTCGAGAGGCGCCAGCATCTGCAAAAGCTCC 465
QY      833 CGGCGCGCGCGAGCGCTGGATACCAACAGTAGTACCCATACGACGTGCCAGACTACGCACTT 892
Db      464 CGGCGCGCGCGA----- 453
QY      893 CTGCGCGCTGGATACCAACTATGCTTCAGCTCCACGAGAGAACTGCTGCGTGGCGGAG 952
Db      452 ---GCMCTGGACACCAACTATGCTTCAGCTCCACGAGAGAACTGCTGCGTGGCGGAG 396
QY      953 CTCTACATTTCTGCTGGGCGCTGCTGGAGTGGATTCATGAACCCCAAGGCGCTAC 1012
Db      395 CTGTACATTTACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGCGTAC 336
QY      1013 CATGCCAATTTCTGCTGGGCGCTGCTGCTACATCTGGAGCCTAGACACTAGTACAGC 1072
Db      335 CATGCCGACTTCTGCTGGGCGCTGCTGCTACATTTGGAGCCTGGACACGCAAGTACAGC 276
QY      1073 AAGTCTGCTGGCTCTGTATACAAACAGCACAAACCGCGCGCTGCGGCGCGCTGCTGCTG 1132
Db      275 AAGTCTGCTGGCGCTGTACAAACAGCATAAACCGCGCGCTGCGGCGCGCTG- TGGTG 217
QY      1133 CGGAGGCGCTGGAGCGCACTGCCATCGTGTACTACGTGGGCGCGCAAGCCCAAGTGGAG 1192
Db      216 CGGAGGCGCTGGAGCGCGCTGCCATCGTGTACTACGTGGGCKKCAAGCCCAAGTGGAG 157
QY      1193 CAGCTGTCCAAACATGATGCTGCTGCTGCAAGTGCAGCTGAGGCGCGCGCGCGCCACCA 1252
Db      156 CAGCTGTCCNACATGATGCTGCTGCTGCAAGTGCAGCTGAGATCCGCGCGCGCGCGC 97
QY      1253 GCCCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1303
Db      96 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 37
QY      1304 TGTATTTAAGCA-CATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1338
Db      36 TGTATTTAAGNACCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1

```

```

RESULT 9
AL530081
LOCUS      859 bp      mRNA      linear      EST 23-MAY-2003
DEFINITION AL530081 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
            CDNA clone CS0DD009YM0f 5-PRIME, mRNA sequence.
ACCESSION  AL530081
VERSION     AL530081.2 GI:31067916
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 859)
            Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
            Full-length cDNA libraries and normalization
            Unpublished
            On Feb 13, 2001 this sequence version replaced gi:12793574.
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            was normalized. library was constructed by Life technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            9160.r For more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DD009BG030p1&cluster=9160.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DD009BG03Qp1.
            Location/Qualifiers
FEATURES

```

```

source
1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD009YM0f"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT: 179 a 272 c 257 g 148 t 3 others
ORIGIN
Query Match      42.4%; Score 577.4; DB 9; Length 859;
Best Local Similarity 87.0%; Pred. No. 8e-116;
Matches 643; Conservative 2; Mismatches 93; Indels 1; Gaps 1;
QY      6 CCGAGATGGCGCTTCGGGGCTCGGGCTCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65
Db      121 CCCCCATGCCCGCTTCGGGGCTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY      66 TAGTCTGACGCTGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125
Db      181 TGGTCTGACGCTGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY      126 AGCTGCTGAAGCGGAGCGCATCGAGCGCATTCGCGCGCGAGATTCCTCCAAAGCTTCGCG 185
Db      241 AGCTGCTGAAGCGGAGCGCATCGAGCGCATTCGCGCGCGAGATTCCTCCAAAGCTTCGCG 300
QY      186 TTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 245
Db      301 TGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY      246 CTCCTTTACAACAGTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 305
Db      361 CCGCTGTACAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY      306 CAGAGCGCGCGCTACTACCGCAAGGAGGTACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 365
Db      421 CTGAGCGCGCGCTACTACCGCAAGGAGGTACCGCGCGCTGCTGCTGCTGCTGCTGCTGCT 480
QY      366 AANTCTATGATTAATTCGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 425
Db      481 AANTCTATGACAAGTTCGAAGCAGAGTACACACAGCATATATATATGTTCTTCAACACATCAG 540
QY      426 AGCTCGGGAAGCGCTGCGCGAAACCTGTATTGCTCTCTCGCGCGAGAGTGGCGCGCTGCTGA 485
Db      541 AGCTCCGAGAAGCGCTACCTGAACCCGTTGCTCTCCCGGCGAGAGTGGCTGCTGCTGA 600
QY      486 GGCTCAAGTTAAAGTGGAGCAGCAGCTGGAGCTATACCGAAATACAGCAATGATTCCT 545
Db      601 GGCTCAAGTTAAAGTGGAGCAGCAGCTGGAGCTATACCGAAATACAGCAATGATTCCT 660
QY      546 GCGCTTACTCTAGCAACCGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 604
Db      661 GCGGATACCTCAGCAACCGGCTGCTGGCAGCCAGCGAGTCCCGAGAGTGTGTATCTTTT 720
QY      605 GATGTACCGGAGTTGTGCGGCGAGTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 664
Db      721 GATGTACCGGAGTTGTGCGGCGAGTGGTGGAGGAGAAATGAGGCGCTTCGCG 780
QY      665 CTCAGTGGCGCGCTCTCTCTGACAGCAAGATACACACTCCAGCTGGGAATTAACGGG 724
Db      781 CTAGCGCGCGCTCTCTCTGTCAGCAGCGGATAACACACTGCAAGTGRACATCAACGGG 840
QY      725 TTCAATTCCTGGCGCGCGCGG 743
Db      841 TTCACCTACCGCGCGCGCGAG 859

```

RESULT 10
BE312000

```

LOCUS      BE312000      713 bp      mRNA      linear      EST 26-OCT-2000
DEFINITION 601154768F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3510592 5',
mRNA sequence.
ACCESSION   BE312000
VERSION     BE312000.1 GI:9130123
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 713)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL     NIH-MGC http://mgc.nci.nih.gov/.
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
            Plate: LLCMI95 row: h column: 17
            High quality sequence start: 2
            High quality sequence stop: 713.

FEATURES             source
     source
     1..713
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:3510592"
     /tissue_type="neuroblastoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_19"
     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
     EcoRI; cDNA made by oligo-dr priming. Directionally
     cloned into EcoRI/XhoI sites using the following 5'
     adaptor: GGCAACGAG(G). Library constructed by Ling Hong
     in the laboratory of Gerald M. Rubin (University of
     California, Berkeley) using ZAP-cDNA synthesis kit
     (Stratagene) and Superscript II RT (Life Technologies).
     Note: this is a NIH_MGC Library."
     157 a 226 c 211 g 119 t

BASE COUNT  157 a 226 c 211 g 119 t
ORIGIN
Query Match      41.9%; Score 570.6; DB 10; Length 713;
Best Local Similarity 87.5%; Pred. No. 2.3e-114;
Matches 624; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 135 AGCGGAAGCGATCGAGGCCATTGCGGGCCAGATTCTGTCCAAAGCTTGGCTTGCAGCC 194
Db 1 AGCGGAAGCGATCGAGGCCATTGCGGGCCAGATTCTGTCCAAAGCTTGGCTTGCAGCC 60

Qy 195 CCCGAGCCAGGGGAGCTCCGCGCGCGCGCTGCTGAGGCGAGTACTGGCTTTACA 254
Db 61 CCCGAGCCAGGGGAGGTGTCGCGCGCGCGCGCTGCTGAGGCGCGCTGCTGAGGCG 120

Qy 255 ACAGTACCCGCGACCGGCTGCGCGGGGAAAGTGTGCAACCCGAGCCCGAGAGCGCG 314
Db 121 ACAGCACCAGCGGCGGCTGCGCGGGGAGAGTGCAGAACCGGAGCGCGCTGAGCGCG 180

Qy 315 ACTACTAGCCCAAGAGGTGTCACCGCGTGTATGTTGGTGAAGGGGCAACCAATCTATG 374
Db 181 ACTACTAGCCCAAGAGGTGTCACCGCGTGTATGTTGGTGAAGGGGCAACCAATCTATG 240

Qy 375 ATAATTCAGGCGACCCCGACAGCTTATATGCTGTTCACACGCTCGGAGCTCCGGG 434
Db 241 ACAAGTTCAACGAGGATACACAGCATATATATGTTCTTCACACATCAGAGCTCCGAG 300

Qy 435 AAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTCGCGCTGCTGAGGCTCAAGT 494
Db 301 AAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCGAGAGCTCGGCTCTCTGAGGCTCAAGT 360

```

```

Qy 495 TAAAGTGGAGCAGCAGCTGGAGCTATACACAGAAATACAGCAATATGTTCTTGGCGTACC 554
Db 361 TAAAGTGGAGCAGCAGCTGGAGCTATACACAGAAATACAGCAATATGTTCTTGGCGTACC 420

Qy 555 TCAGCAACCGGCTGCTGGCGCGCGCAGTCACTACCGGAGTGGCTGCTCTTGTGATGTCACGG 614
Db 421 TCAGCAACCGGCTGCTGGCGCGCGCAGTCACTACCGGAGTGGCTGCTCTTGTGATGTCACGG 480

Qy 615 GAGTTGTGGCGCAGTGGCTGACCCGAGAGAGGCTATACAGGGTGTTCGCGCTCAGTGGCC 674
Db 481 GAGTTGTGGCGCAGTGGCTGACCCGAGAGAGGCTATACAGGGTGTTCGCGCTCAGTGGCC 540

Qy 675 ACTTCTCTGTGACGCAAAAGATAACACACTCCAGCTGGGAAATACAGGGTGTTCATCTGT 734
Db 541 ACTTCTCTGTGACGCAAAAGATAACACACTCCAGCTGGGAAATACAGGGTGTTCATCTGT 600

Qy 735 GCCCGCGGGTGACCTGGCCACCATTCACGGCATGAAACCGCGCTTCTGCTCTCTCATGG 794
Db 601 GCCCGCGGGTGACCTGGCCACCATTCACGGCATGAAACCGCGCTTCTGCTCTCTCATGG 660

Qy 795 CCACCCCGTGGAGAGGGCCAGCAGCTGCACAGCTCCCGGACCCCGCGAGCC 847
Db 661 CCACCCCGTGGAGAGGGCCAGCAGCTGCACAGCTCCCGGACCCCGCGAGCC 713

RESULT 11
LOCUS      BE260971      717 bp      mRNA      linear      EST 26-OCT-2000
DEFINITION 601153715F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3509931 5',
mRNA sequence.
ACCESSION   BE260971
VERSION     BE260971.1 GI:9132709
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 717)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL     NIH-MGC http://mgc.nci.nih.gov/.
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
            Plate: LLCMI93 row: m column: 04
            High quality sequence stop: 713.

FEATURES             Location/Qualifiers
     source
     1..717
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:3509931"
     /tissue_type="neuroblastoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_19"
     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
     EcoRI; cDNA made by oligo-dr priming. Directionally
     cloned into EcoRI/XhoI sites using the following 5'
     adaptor: GGCAACGAG(G). Library constructed by Ling Hong
     in the laboratory of Gerald M. Rubin (University of
     California, Berkeley) using ZAP-cDNA synthesis kit
     (Stratagene) and Superscript II RT (Life Technologies).
     Note: this is a NIH_MGC Library."
     158 a 226 c 213 g 120 t

BASE COUNT  158 a 226 c 213 g 120 t
ORIGIN
Query Match      41.3%; Score 562.6; DB 10; Length 717;
Best Local Similarity 87.4%; Pred. No. 1.3e-112;

```

Matches 627; Conservative 0; Mismatches 89; Indels 1; Gaps 1;

QY 136 GCGAAGCGATCAGGCCATCGCGGCCAGATCTGTCCAAAGTTCGGCTTGCCAGCC 195
 Db 1 GCGAAGCGATCAGGCCATCGCGGCCAGATCTGTCCAAAGTTCGGCTTGCCAGCC 60

QY 196 CCCGAGCGAGGGGACGCTGCGCGCGCGCGCTGCTGAGGCGAGTACTGGCTTTACAA 255
 Db 61 CCCGAGCGAGG-GGAGGTGCGCGCGCGCGCTGCGCGCGCGCTGCTGAGGCGAGT 119

QY 256 CAGTACCGCGACCGGTAGTCGGGGAAAGTGTGGAACCGGAGCGCGAGCGAGCGGA 315
 Db 120 CAGCACCAGCGACCGGTGGCGCGGGAGAGTGCAGAACCGGAGCGCGAGCGAGCGGA 179

QY 316 CTACTAGCGCAAGGAGGTCAACCGCGTGTAAATGGTGGAAAGCGGCAACCAATCTATGA 375
 Db 180 CTACTAGCGCAAGGAGGTCAACCGCGTGTAAATGGTGGAAAGCGGCAACCAATCTATGA 239

QY 376 TAAATTCAGGGGACCGCCGACAGCTTATATATGCTTCAACACGCTCGGAGCTCCGGA 435
 Db 240 CAGTTCAGGAGAGTACACAGCATATATATGCTTCAACACATCAGAGCTCCGAGA 299

QY 436 AGCGGTGCGGAACTGTATGCTCTCGGGCAGAGCTCGGCTGCTGAGGCTCAAGTT 495
 Db 300 AGCGGTACTCAACCGGTGTGCTCTCGGGCAGAGCTCGGCTGCTGAGGCTCAAGTT 359

QY 496 AAAAGTGGAGCAGCGTGGAGTATACCAAGATACAGCAATGATTCCTGGCGCTACCT 555
 Db 360 AAAAGTGGAGCAGCGTGGAGTATACCAAGATACAGCAATGATTCCTGGCGCTACCT 419

QY 556 CAGCAACCGGTGCTGGCGCCGACGTGACTCACCGAGTGTCTTTCATGCTCAACCGG 615
 Db 420 CAGCAACCGGTGCTGGCGCCGACGTGACTCACCGAGTGTCTTTCATGCTCAACCGG 479

QY 616 AGTGTGGCGAGTGGCTGACCCGACAGAGGCTATAGAGGTTTCGCCCTCAGTGCCTCA 675
 Db 480 AGTGTGGCGAGTGGTGTGACCGGTGGAGGGAATTTGAGGGCTTTGCGCTTACGCCCA 539

QY 676 CTCTTCCTCTGACAGCAAGATACACACTCCAGTGTGAATTAACGGGTTCATCTGG 735
 Db 540 CTGCTCTGTGACAGCAGGATACACACTGCAAGTGGACATCAACGGGTTCATCAGCG 599

QY 736 CGCGCGGCTGACCTGGCCACCTTCAGGCGATCAACCGCGCTTCTGCTCCTCATGGC 795
 Db 600 CGCGCGAGTGCCTGGCCACCTTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 659

QY 796 CACCCGCTGAGAGGGCCAGCACCTGACAGCTCCGCGCAGCGCGAGCGCTGGA 852
 Db 660 CACCCGCTGAGAGGGCCAGCACCTGACAGCTCCGCGCAGCGCGAGCGCTGGA 716

RESULT 12
 BQ675698
 LOCUS
 DEFINITION BQ675698 902 bp mRNA linear EST 15-JUL-2002
 AGENCOURT_8036532 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6211917
 5' mRNA sequence.

ACCESSION BQ675698
 VERSION BQ675698.1 GI:21786532
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 902)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LCM2375 row: c column: 22
 High quality sequence stop: 599.

FEATURES
 source
 1..902
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6211917"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_102"

Site 1: XhoI;
 Note: Ordan: salivary gland; Vector: pOR7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 194 a 292 c 255 g 161 t

Query Match 41.3%; Score 562; DR 13; Length 902;
 Best Local Similarity 86.7%; Pred. No. 1.9e-112;
 Matches 619; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 156 ITCCGGCGCAGATCTGTCCAAAGCTTGGGCTTGCCAGCCCGCGAGCGAGGGAGCTGC 215
 Db 1 TCCGGCGCAGATCTGTCCAAAGCTTGGGCTTGCCAGCCCGCGAGCGAGGGAGCTGC 60

QY 216 CGCCGGCGCCGCTGCTGAGGCGAGTACTGGCTTTACACAGTACCCGCGAGGGGTAG 275
 Db 61 CGCCGGCGCCGCTGCTGAGGCGAGTACTGGCTTTACACAGTACCCGCGAGGGGTAG 120

QY 276 CGGGGAAAGTGTGCAACCGCGAGCGCGAGCGAGGGGAGTACTAGCCCAAGAGGTCA 335
 Db 121 CGGGGAGAGTGTGCAACCGCGAGCGCGAGCGAGCGAGTACTAGCCCAAGAGGTCA 180

QY 336 CGCGCTGCTAATGCTGGAAGCGGCAACCAATCTATGATAAATTCAGGCGACCCGCC 395
 Db 181 CGCGCTGCTAATGCTGGAAGCGGCAACCAATCTATGATAAATTCAGGCGAGTACAC 240

QY 396 ACAGCTTATATGCTGTCAACAGCTCGGAGCTCGGGAGAGCGGTGCGCGAAGCTGAT 455
 Db 241 ACAGCTTATATGCTGTCAACAGCTCGGAGAGCTCGGAGAGCGGTGCTGAAACCGTGT 300

QY 456 TGCTCTCTCGGCGAGAGTGGCGCTGTGAGGCTCAAGTTAAAGTGGAGCAGCAGCTGG 515
 Db 301 TGCTCTCTCGGCGAGAGTGGCGCTGTGAGGCTCAAGTTAAAGTGGAGCAGCAGCTGG 360

QY 516 AGCTATACAGAAATACAGCAATGATTCCTGCGGTACTCTCAGCAACCGCGCTGCGCCC 575
 Db 361 AGCTATACAGAAATACAGCAATGATTCCTGCGGTACTCTCAGCAACCGCGCTGCGCCC 420

QY 576 CCAGTGACTCACCGGAGTGGCTGTCTTTGATGTACCGGAGTGTGCGCAGTGGCTGA 635
 Db 421 CCAGTGACTCACCGGAGTGGCTGTCTTTGATGTACCGGAGTGTGCGCAGTGGCTGA 480

QY 636 CGCGCAGAGAGCTATAGAGGTTTTCGCCCTCAGTGGCCACTCTTCCTCTCAGCAAGAG 695
 Db 481 CGCGTGGAGGGAATTTAGGGCTTTCGCCCTCAGTGGCCACTCTTCCTCTCAGCAAGAG 540

QY 696 ATAACACACTCCACCTCGGAATTAACGGGTTCATGCTGCGCGCGGGGTGACCTGGCCA 755
 Db 541 ATAACACACTCGCAAGTGCACATCAACGGGTTCATGCTGCGCGCGGGGTGACCTGGCCA 600

QY 756 CCATTACGCGCATGAACCGGCGCTTCTCTGCTCTCATGGCCACCCCGCTGGAGAGGCC 815
 Db 601 CCATTACGCGCATGAACCGGCGCTTCTCTGCTCTCATGGCCACCCCGCTGGAGAGGCC 660

```

QY      816 AGCACTGACAGCTCCCG3CACCGCGAGCCCTGGATACCAACAGCTACCAT 869
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      661 AGCATCTGCAAGCTCCCG3CACCGCGAGCCCTGGGACACCATATTGCTT 714

RESULT 13
B1818841
LOCUS      603037307F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5178433 5',
DEFINITION mRNA sequence.
ACCESSION B1818841
VERSION   B1818841.1 GI:15930391
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 925)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL   NIH-MGC http://mgc.nci.nih.gov/
COMMENT   National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished
          Contact: Robert Strauberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L14M11444 row n column: 02
          High quality sequence stop: 874.
          Location/Qualifiers
            1..925
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:5178433"
              /lab_host="DH10B"
              /clone_lib="NIH_MGC_115"
              /note="Organ: pooled brain, lung, testis; Vector:
              pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
              source anonymous pool of 6 male brains, age range 23-27; 1
              male lung, age 27; and 1 male testis, age 69. Library is
              oligo-dT primed and directionally cloned (EcoRV site is
              destroyed upon cloning). Average insert size 1.8 kb,
              insert size range 1-3 kb. Library is normalized and
              enriched for full-length clones and was constructed by C.
              Gruber (Invitrogen). Research Genetics tracking code
              021. Note: this is a NIH_MGC Library."
            170 a 316 c 278 g 161 t

BASE COUNT 170 a 316 c 278 g 161 t
ORIGIN

Query Match 39.3%; Score 534.4; DB 12; Length 925;
Best Local Similarity 87.2%; Pred. No. 2e-106;
Matches 609; Conservative 6; Mismatches 86; Indels 3; Gaps 2;

QY      6 CCGAGATGGCGCCTTCGGGCTGTCGGCTCTGTGGCTGCTGCTGCTGCTG 65
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      146 CCCCATACCCCGCTCCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTAC 205

QY      66 TAGTGCTGACGCTGGCGCGCGCGGCTGTCACCTGCAGACCATCCAGATGG 125
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      206 TGGTGTGACGCTTGGCGCGCGCGGCTGTCACCTGCAGACCATCCAGATGG 265

QY      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAACTTCGGC 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      266 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAACTTCGGC 325

QY      186 TTGCCAGCCCCCGAGCAGGGGACGTGCGCGCGCGCGCTGCTGCTGAGGCGAGTACTGG 245
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      326 TCGCAGCCCCCGAGCAGGGGAGGTGCGCGCGCGCGCTGCTGCTGAGGCGAGTACTGG 385

```



```
BASE COUNT      150 a      240 c      283 g      153 t      15 others
ORIGIN

Query Match      38.3%; Score 521.4; DB 9; Length 841;
Best Local Similarity 80.1%; Pred. No. 1.4e-103;
Matches 687; Conservative 8; Mismatches 106; Indels 57; Gaps 5;

QY 438 CGGTGCGGAACCTGTATTGCTCT-CTCGGCAGAGTGGCGCTGTGAGGCTCAAGTTA 496
DB 841 CGGTACCTGMAACCTTGTGTCTCCCGGCAGAGTGTCTGTCTAAGGCTCAAGTTA 782
QY 497 AAGTGGAGCAGCAGTGGAGCTATACAGAAATACAGAAATATTCCTGGCGCTACCTC 556
DB 781 AAGTGGAGCAGCAGTGGAGCTGTACAGAAATACAGAAATATTCCTGGCGCTACCTC 722
QY 557 AGCAACCGGCTGCTGGCGCCCTAGTACTCACCGAGTGGCTCTTTGATGTACCGGA 616
DB 721 AGCAACCGGCTGCTGGCGCCCTAGTACTCACCGAGTGGCTCTTTGATGTACCGGA 662
QY 617 GTTGTGCGGAGTGGCTGAC-CGCGAGAGAGCTATAGAGGGTTCGCGCTCAGTGGCCAC 676
DB 661 GTTGTGCGGAGTGGCTGAC-CGCGAGAGAGTTCAGG--TTTCGCTTAGCGCCNAC 604
QY 677 TCTTCTCTGACCAAGAATAACACACTCCAGCTGGAAATTAACGGTTCAATTCGGC 736
DB 603 GGCTCTCTTAACAGCAGGAGTAACACACTGCAAGTGAACATCAACGGGTTAACTACCGG 544
QY 737 CGCCGGGTGAGCTGGCCACATTCACCGGCATGAACCGGCTTCCTGCTCTCATGGCC 796
DB 543 CGCCGAGGTGAGCTGGCCACATTCCTGGCATGAACCGGCTTCCTGCTCTCATGGCC 484
QY 797 ACCCGCTGGAGAGGCGCCAGCAGCTCCAGAGTCCCGGAGCGCGAGCGCTGGATACC 856
DB 483 ACCCGCTGGAGAGGCGCCAGCTTCGAAAGCTCCCGGAGCGCGA----- 436
QY 857 AACAGCTACCCATAGAGCTGTCAGACTACGATCTCTGGCGCTCGATACCAACTACTGC 916
DB 435 -----GCCCTGGACCACTATTGC 415
QY 917 TTCAGCTCCACGAGAGAAATGTGCTGCTGGGAGCTCTACATGTCTCCGGAAGGAC 976
DB 414 TTCAGCTCCACGAGAGAAATGTGCTGCTGGGAGCTGTACATTAAGCTTCGGAAGGAC 355
QY 977 CTGGGCTGAAGTGGATCA-GAAGCCAGAGGCTACCATGCTCAATTTCTGCTGGGCGCC 1036
DB 354 CTCGGCTGGAGTGCATCCAGAGCCAGAGGCTWCCATGCTGCTCGGCGCC 295
QY 1037 TGTGCTTACATCTGAGCGCTAGACACTCAGTACAGCAAGGCTCTGGGCTCTGTACACAG 1096
DB 294 TGCCTCTACATTTGAGGCTTGACACGAGGAGCAAGAGTCTGGGCTCTGTACACAG 235
QY 1097 CACAACCGGCGCTGCGGCGCGCTGCTGCTGCGCGAGGCGCTGGAGCCACTGCCC 1156
DB 234 CATACCCCG--GCCTGGCGCGCGCTGCTGCTGCGCGAGGCTGGAGCGCTGCCC 178
QY 1157 ATGCTGTACTAGTGGCGCGAAGCCCAAGGTGAGCAGCTGTCCAAATGATGCTGGCT 1216
DB 177 TTCGCTTACTAGTGGCGCGAAGCCCAAGGTGAGGCGGCTGTCCAAATGATGCTGGCTG 118
QY 1217 TCTGCAAGTGTAGGCTGAGGCTGCGCGCGCGCCAGCGCGCGCGCGCGCGCGCGCG 1276
DB 117 TCTGCAAGGCGCTGAGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 58
QY 1277 CCGACCGCGCGCGCGCT 1294
DB 57 CCGCGCGCGCGCGCGCT 40
```

```
RESULT 15
BI084718/c
LOCUS
DEFINITION 602869722T1 NIH_MGC_102 Homo sapiens CDNA clone IMAGE:5014421 3',
          BI084718          956 bp mRNA linear EST 20-JUN-2001
```

```
mrna sequence.
BI084718
BI084718.1 GI:14503048
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1821 row: 1 column: 06
High quality sequence start: 3
High quality sequence stop: 793.
FEATURES
Location/Qualifiers
1..956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5014421"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/Note="Organ: salivary gland; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGCAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using 2AP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 172 a 273 c 312 g 199 t
ORIGIN
Query Match 38.3%; Score 521.4; DB 12; Length 956;
Best Local Similarity 76.8%; Pred. No. 1.4e-103;
Matches 755; Conservative 0; Mismatches 161; Indels 67; Gaps 7;
QY 329 GAGGTACACCGCGCTGCTAATGTGGAAGCGGCACCAATCTATGATAAATTCAGGGC 388
DB 955 GAGGTAAACCGGTGCTAATGTGGAAGCGGCACCAATCTATGACGAGGTCAGCAGA 896
QY 389 ACCCGCCACAGCTTATATATGCTTCAACAGCTC--GGAGCTCCGGGAAGCGTGGCGG 446
DB 895 CTACACACAGCCATAAATATGCTTTCACACATTCAGGAGCTTCGGAACCGGTACCT 836
QY 447 AACCTGTATGCTCTCTCGGCGAGAGCTGGCGCTG----CTGAGGCTCAAGTTAAAGT 501
DB 835 GAACCGTGTGTGCTCTTCCCGGGAAGAGCTGCCGCTGCTGAGGCTCAAGTTAAAGT 776
QY 502 -GGAGCAGCAGCTGGAGCTATACCAAGAAATACAGCAATGATTCCTGGCGCTACCTCAGCA 560
DB 775 GGACCCAGCAGCTGGAGCTGTACCCAGGAATACAGCAATTCCTGGCGATACCTCAGCA 716
QY 561 ACCGCGTGTGCGCCCGCAGTGA--CTCACCGGAGTGGCTGCTCTTTGATGTACCGGAGT 618
DB 715 ACCGCGTGTGCGCCCGCAGCTCGCCAGCACTCGCCAGAGTGGTTATCTTTTGTGTCACCGGAGT 656
QY 619 TGTGCGGAGTGGCTGACCGCCAG--AGAGGCTATAGAGGCTTTTCGCTCAGTGGCCACT 677
DB 655 TGTGCGGAGTGGTGTGAGCGGTGGAAGGGGAATTTGAGGCTTTTCGCTTGTAGGCGCCACT 596
QY 678 CTTCTCTCAGCAGCAAGAAATACACACTCCAGCTCGAAATTAACGGGTTCAATTCGCCC 737
```

[illegible]

Search completed: October 9, 2003, 18:34:02
Job time : 3231 secs